5-2012

Investigations of Pharmacokinetic Challenges in Premature Infants

Yi Zhang
University of Tennessee Health Science Center

Investigations of Pharmacokinetic Challenges in Premature Infants

Abstract

Premature infants (gestational age less than 37 weeks) are considered a vulnerable patient population due to their immaturity at birth. Currently, off-label prescribing is common in younger pediatric populations, especially in premature neonates and infants, which is a primary group receiving intensive care. Unique pharmacokinetic (PK) challenges—such as limited blood volume and frequency of blood sample collections, rapid growth and continuous developmental changes, complexity of pediatric studies as well as scientific, practical, and ethical concerns— lead to the current lack of PK information and empirical dosing in premature neonates and infants. In this research, several approaches were investigated to overcome these PK challenges. We first developed and validated an accurate and sensitive LC-MS/MS method that can simultaneously quantitate multiple drugs frequently used in pediatric pharmacotherapy using a small volume of plasma. Additionally, a modeling and simulation (M&S) approach was explored in the theophylline population pharmacokinetic (PopPK) study in order to get an appropriate study design with the optimized sample size. Finally, PopPK of caffeine was investigated in premature infants using clinical data. Optimized dosing regimens were developed based on the PopPK model and dose-finding simulation study.

Due to the limitation in sample volume, an assay that can simultaneously determine multiple drugs allows for gaining maximal information from PK studies while minimizing the burden of blood collection in pediatric patients. Acetaminophen, caffeine, phenytoin, ranitidine, and theophylline are widely used in the pharmacotherapy of premature and term neonates, but only limited information is currently available on the PK of these medications in premature neonates. An accurate, sensitive and reliable LC-MS/MS assay was developed and validated using 50 µL human plasma specimens to simultaneously quantitate these five drugs with the mean accuracy ranging from 87.5 to 115.0%. The intra-day and inter-day precisions ranges from 2.8% to 11.8%, 4.5% to 13.5% respectively. This assay quantifies a range of 12.2 to 25,000 ng/mL for acetaminophen, phenytoin, and ranitidine, a range of 24.4 to 25,000 ng/mL for theophylline, and a range of 48.8 to 25,000 ng/mL for caffeine. These ranges cover each drug's therapeutically used concentrations in the neonatal group. No significant interference effects from hemolysis, lipemia and hyperbilirubinemia were noted when these factors existed separately or were combined. Additionally, no significant matrix effect was observed for the developed bioanalytical assay.

We then evaluated the impact of sample size on the robustness of PopPK parameter estimates in observational studies in premature neonates using a simulation approach with theophylline as the model drug. Simulated datasets for each sample size (9–200 subjects per study) with a mixed and unbalanced sampling design were first generated with the incorporation of changes in birth weight, body weight, and postnatal age (PNA) in premature neonates. The median PopPK parameters for theophylline estimated from the simulated datasets were generally in close agreement with those of the originating model across all tested sample sizes. While the accuracy, precision and power to parameter estimation benefit from increases in the number of subjects included in the study, an observational study designs with < 20 premature neonates and unbalanced sampling are inadequate to allow for the precise estimation of theophylline PopPK parameters. Furthermore, the results indicate that the impact of sample size on the power of the study was deeply influenced by the parameter of interest and the selected precision level. To detect all three covariate effects studied in this research with a power > 0.8, a sample size of 20, 40 and 60 subjects is required to reach the significant level of P = 0.05, P = 0.01 and P = 0.001, respectively. The application of PopPK modeling and simulation provides a useful approach to estimate the number of subjects needed to confidently detect the potential covariate effects on PK parameters under a specific sampling strategy—randomized and unbalanced blood sampling schedules, which is consistent with actual pediatric clinical settings.

Apnea of prematurity (AOP) is one of the major concerns in premature neonates. Caffeine is currently the first-line pharmacotherapy frequently used for the treatment of AOP. A PopPK model of caffeine was developed in premature neonates, and potential sources of variability of PK behavior for caffeine were also identified. A one-compartment model was chosen to describe the PK characteristics of caffeine in premature infants, covering a gestational range of 23 to 31 weeks with an age of up to 116 days. Body weight (WT), postconceptional age (PCA) and a low gestational age (GA) of < 25 weeks were found to be important predictors explaining the between-subject variability of caffeine PK in premature infants receiving caffeine treatment. The typical patient in the studied premature neonate population, i.e., a patient with WT of 1.5 kg, PCA of 32 weeks and with a GA > 25 weeks, is estimated to have a CL of 0.0164 L/hr and a V of 0.94 L. We also investigated the application of this PK knowledge to facilitate the development of optimal dosing regimens further through simulation, particularly to correlate steady state concentrations with response at the different dosing regimens for various age/body size groups using trial simulation. A dosing interval of 24 hours is shown to be successful with respect to the proposed target concentrations in all simulated groups. With the proposed dosing regimens, the predetermined target was attained and the simulated median trough plasma concentrations were between 8 and 20 mg/L throughout the treatment period. The dose-finding simulations based on the developed PopPK model may provide more benefit while allowing the clinicians to compare various dosing regimens and bridge the plasma caffeine levels with responses at different PCAs and different WTs.

In summary, different approaches were investigated in this study to overcome the unique PK challenges in the premature neonates and infants. A full model-based simulation approach was developed to determine an optimal sample size for PopPK study in premature neonates with the consideration of changes in birth weight, body weight, and PNA. In addition, a PopPK model was developed for caffeine in premature infants and optimal dosing regimens were proposed to reach the therapeutic target concentrations rapidly based on the PopPK model. Together with the developed LC-MS/MS assay, which is highly sensitive, accurate and reliable, population-based modeling and simulation are highly useful in supporting clinical PK studies in premature neonates and infants.

Document Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Program
Pharmaceutical Sciences

Research Advisor
Bernd Meibohm, Ph.D.

Keywords
premature infants, LC-MS/MS, modeling and simulation, population pharmacokinetics, sample size estimation, dosing optimization

Subject Categories
Medicine and Health Sciences | Pharmaceutics and Drug Design | Pharmacy and Pharmaceutical Sciences

INVESTIGATIONS OF PHARMACOKINETIC CHALLENGES IN PREMATURE INFANTS

A Dissertation
Presented for
The Graduate Studies Council
The University of Tennessee
Health Science Center

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy
From The University of Tennessee

By
Yi Zhang
May 2012

To

my husband Shen Li,

my daughter Jiayun Lily Li, my son David Zelin Li,

my mother Jingzun Wang and my father Hongrui Zhang,

for

their never-ending love and support.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest and sincere gratitude to my mentor and advisor, Dr. Bernd Meibohm, for providing me with the opportunity to work in his lab on this project. A tremendous thank you to him for introducing the world of pharmacokinetics and quantitative pharmacology to me. Without his professional guidance, advice, support, persistent encouragement and patience over the years, my dissertation would not have been accomplished. He instilled in me his knowledge and his enthusiasm for science. The training in his lab has greatly benefited me in my understanding of science and my personal life.

I also would like to express sincere appreciation to the other members of my committee, Dr. Michael Christensen, Dr. John Carl Panetta, Dr. Fridtjof Thomas and Dr. Charles Ryan Yates for their invaluable suggestions, direction and assistance throughout the learning and writing process.

A special thank you goes to Dr. Anthony Christensen for providing me with the caffeine dataset, without which a major portion of my dissertation work would not have been possible.

I would like to express my gratitude to my friend Dr. Yanhua Qu for helping me develop a Perl script that considerably speeded up the portion of my work in Chapter 3 project.

I would like especially to acknowledge Dr. Charles Ryan Yates for providing me with the opportunity of working as a research assistant in his lab to establish my lab skills under his guidance and encouragement before I enrolled in the graduate program.

A special thank you to all my friends, former and current colleagues: Margaret Thomson, Wararat Limothai, Dora Babu Madhura, Ashit R. Trivedi, Chetan Rathi, Josiah Ryman, Wenhui Zhang, Manish Gupta, Lisa Tang, Nageshwar R. Budha, Nathanial Dirks, Nitin Mehrotra, Satyendra Suryawanshi, Pavan Vaddady, Lei Diao, Pengfei Song, Les Stuart, Fei Ma, Karin Thompson, Chaela Presley and Jordan J. Toutounchian for their encouragement, helpful discussions and wonderful company over all these years. In particular, I would like to thank Dr. Nageshwar R. Budha for his support and helpful suggestions on the LC-MS/MS experiments. In such a united family, I received support any time I needed it. This environment made it enjoyable and inspiring to work every day.

I'm also grateful to the Department of Pharmaceutical Sciences, The University of Tennessee Health Science Center. I cherish the nine years spent here. My thanks are extended to the entire faculty and staff who assisted me in so many ways over the past years. A special thank you goes to Brenda Thornton, Benita Williams, Cynthia Crowe, Corliss Finlay and former staff members Faith B. Barcroft, Julia Davis, Felicia Martin for their suggestions and administrative help.

Finally and most importantly, I would like to thank all my family members for their love and support. First special thanks to my mother for taking such good care of me from heaven during all of my endeavors; without her love and watchful eyes, I could not have done this. I am eternally thankful to my parents and my brother Ming. I'm also thankful to my husband Shen and my children Lily and David. Their unconditional love and belief in me have been major driving forces to keep me moving forward with my graduate education. No words would be enough to thank them for their encouragement, support, and patience during the completion of this dissertation. Thank you, Shen, for waiting six years for me to complete the graduate study. You always support anything I attempt to accomplish. I notice all the big sacrifices you make and all the extra little things you do for me and for our family. Life is so wonderful just because I can share it with you. In retrospect, our family has been tested and blessed, and I would not trade this precious experience for anything. Thank you, and God be with us!

ABSTRACT

Premature infants (gestational age less than 37 weeks) are considered a vulnerable patient population due to their immaturity at birth. Currently, off-label prescribing is common in younger pediatric populations, especially in premature neonates and infants, which is a primary group receiving intensive care. Unique pharmacokinetic (PK) challenges—such as limited blood volume and frequency of blood sample collections, rapid growth and continuous developmental changes, complexity of pediatric studies as well as scientific, practical, and ethical concerns—lead to the current lack of PK information and empirical dosing in premature neonates and infants. In this research, several approaches were investigated to overcome these PK challenges. We first developed and validated an accurate and sensitive LC-MS/MS method that can simultaneously quantitate multiple drugs frequently used in pediatric pharmacotherapy using a small volume of plasma. Additionally, a modeling and simulation (M&S) approach was explored in the theophylline population pharmacokinetic (PopPK) study in order to get an appropriate study design with the optimized sample size. Finally, PopPK of caffeine was investigated in premature infants using clinical data. Optimized dosing regimens were developed based on the PopPK model and dose-finding simulation study.

Due to the limitation in sample volume, an assay that can simultaneously determine multiple drugs allows for gaining maximal information from PK studies while minimizing the burden of blood collection in pediatric patients. Acetaminophen, caffeine, phenytoin, ranitidine, and theophylline are widely used in the pharmacotherapy of premature and term neonates, but only limited information is currently available on the PK of these medications in premature neonates. An accurate, sensitive and reliable LC-MS/MS assay was developed and validated using 50 µL human plasma specimens to simultaneously quantitate these five drugs with the mean accuracy ranging from 87.5 to 115.0%. The intra-day and inter-day precisions ranges from 2.8% to 11.8%, 4.5% to 13.5% respectively. This assay quantifies a range of 12.2 to 25,000 ng/mL for acetaminophen, phenytoin, and ranitidine, a range of 24.4 to 25,000 ng/mL for theophylline, and a range of 48.8 to 25,000 ng/mL for caffeine. These ranges cover each drug's therapeutically used concentrations in the neonatal group. No significant interference effects from hemolysis, lipemia and hyperbilirubinemia were noted when these factors existed separately or were combined. Additionally, no significant matrix effect was observed for the developed bioanalytical assay.

We then evaluated the impact of sample size on the robustness of PopPK parameter estimates in observational studies in premature neonates using a simulation approach with theophylline as the model drug. Simulated datasets for each sample size (9-200 subjects per study) with a mixed and unbalanced sampling design were first generated with the incorporation of changes in birth weight, body weight, and postnatal age (PNA) in premature neonates. The median PopPK parameters for theophylline estimated from the simulated datasets were generally in close agreement with those of the originating model across all tested sample sizes. While the accuracy, precision and power to parameter estimation benefit from increases in the number of subjects included in the

study, our simulation showed observational study designs with < 20 premature neonates and unbalanced sampling are inadequate to allow for the precise estimation of theophylline PopPK parameters. Furthermore, the results indicate that the impact of sample size on the power of the study was deeply influenced by the parameter of interest and the selected precision level. To detect all three covariate effects studied in this research with a power > 0.8, a sample size of 20, 40 and 60 subjects is required to reach the significance level of P = 0.05, P = 0.01 and P = 0.001, respectively. The application of PopPK modeling and simulation provides a useful approach to estimate the number of subjects needed to confidently detect the potential covariate effects on PK parameters under a specific sampling strategy—randomized and unbalanced blood sampling schedules, which is consistent with the actual pediatric clinical settings.

Apnea of prematurity (AOP) is one of the major concerns in premature neonates. Caffeine is currently the first-line pharmacotherapy frequently used for the treatment of AOP. A PopPK model of caffeine was developed in AOP patients, and the potential sources of variability of PK behavior for caffeine were also identified. A one-compartment model was chosen to describe the PK characteristics of caffeine in premature infants, covering a gestational range of 23 to 31 weeks with an age of up to 116 days. Body weight (WT), postconceptional age (PCA) and a low gestational age (GA) of < 25 weeks were found to be important predictors explaining the between-subject variability of caffeine PK in premature infants. The typical patient in the studied premature neonate population, i.e., a patient with WT of 1.5 kg, PCA of 32 weeks and with a GA > 25 weeks, is estimated to have a clearance of 0.0164 L/hr and a volume of distribution of 0.94 L. We further investigated the application of this PK knowledge to facilitate the development of optimal dosing regimens through simulations, particularly to correlate steady state concentrations with response at the different dosing regimens for various age/body size groups using trial simulation. A dosing interval of 24 hours is shown to be successful with respect to the proposed target concentrations in all simulated groups. With the proposed dosing regimens, the predetermined target was attained and the simulated median trough plasma concentrations were between 8 and 20 mg/L throughout the treatment period. The age-specific dose-finding simulations based on the developed PopPK model may provide more therapeutic benefit while allowing the clinicians to compare various dosing regimens and bridge the plasma caffeine levels with responses at different PCAs and different WTs.

In summary, different approaches were investigated in this study to overcome the unique PK challenges in premature neonates and infants. A full model-based simulation approach was developed to determine an optimal sample size for PopPK study in premature neonates with the consideration of changes in birth weight, body weight, and PNA. In addition, a PopPK model was developed for caffeine in premature infants and optimal dosing regimens were proposed to reach the therapeutic target rapidly based on the PopPK model. Together with the developed LC-MS/MS assay, which is sensitive, accurate and reliable, population-based modeling and simulation are highly useful in supporting clinical PK studies in premature neonates and infants.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

θ	Structural model parameter (theta)
ω^2	Between-subject variance estimate (omega-squared)
ABC	ATP-binding cassette family
AOP	Apnea of prematurity
BSV	Between-subject variability
BW	Birth weight
CI(s)	Confidence interval(s)
CL	Clearance
CV	Coefficient of variation
CYP	Cytochrome P450
F (or F1)	Bioavailability
FDA	Food and Drug Administration
FOCE	First order conditional estimation
GA	Gestational age
IS	Internal standard
ka	Absorption rate constant
LC-MS/MS	High performance liquid chromatography with tandem mass spectrometry detection
LGA	Low gestational factor
M&S	Modeling and simulation
NICU	Neonatal intensive care unit
NONMEM	NONMEM software for nonlinear mixed-effects modeling
OFV	Objective function value
PCA	Postconceptional age
PD	Pharmacodynamic(s)
PK	Pharmacokinetic(s)
PNA	Postnatal age
PopPK	Population pharmacokinetic(s)
POSTHOC	Posterior conditional estimation
QC	Quality control
RMSE	Root mean square error
RSD	Relative standard deviation
RSE	Relative standard
TDM	Therapeutic drug monitoring
V	Volume of distribution
VPC	Visual predictive check
WT	Weight

CHAPTER 1. INTRODUCTION

Pediatric Populations and Off-label Medications

Federal legislation and the U.S. Food and Drug Administration (FDA) regulations require that drugs be tested for safety and efficacy in the intended patient population before they are approved for marketing. Prescribing of drugs without appropriate testing and FDA approval for a specific age or diagnosis is thus considered "off-label" or unlicensed use. While more than 200 new drugs have been approved by the FDA over the last decade, the development and evaluation of new medicines in pediatric populations remains limited and particularly scarce in premature and term neonates [1]. The paucity of evidence-based medicines for children is still an extensive, long-standing problem across almost all therapeutic categories including infection, respiratory diseases, central nervous system disorders and gastrointestinal diseases, as well as pain control (**Figure 1-1**) [2-9]. For example, it has been reported that up to 62% of pediatric outpatient visits involve prescribing off-label or unlicensed medications [3]. Among this 62% percent, off-label prescribing was observed in 96% of the pediatric patients with cardiovascular and renal diseases, over 80% of those requiring pain management therapies or diagnosed with gastrointestinal diseases and 67% of patients with pulmonary or dermatologic diseases. Similarly, another study reported that 70% of the medications in pediatric intensive care and 90% of the medications in neonatal intensive care were given in an off-label manner [10].

Pediatric populations are defined as neonates (from birth to 1 month), infants (from 1 month to 2 years), children (from 2 to 12 years), and adolescents (from 12 to < 16 years) in clinical studies. based on the United States Food and Drug Administration (FDA) guidance [11]. Within pediatric populations, premature neonates whose gestational age (GA) is less than 37 weeks are considered the youngest and most vulnerable patient population to adverse drug reactions due to their immaturity at birth. Premature neonates in the intensive care unit have the highest use of off-label medications of any hospitalized patient population. Generally, off-label or empirical drug dosing is not considered unethical or illegal and may in fact be beneficial to the patients. However, given the fact that off-label dosing is usually carried out based upon clinicians' experiences, limited published literature and extrapolations from adult dosing regimens [12], it has been associated with adverse events and raises concerns regarding proper dose selection as well as safety and efficacy [3, 13, 14]. For example, immature liver enzyme activity and low renal excretion function were observed in premature and term neonates compared to adults and older children [15], and these findings may indicate decreased clearance and therefore larger exposure to drugs. A classic example is the "grey baby syndrome" with the administration of chloramphenicol in neonates [16]. The age-related incidence of gray baby syndrome was most found in newborns less than 9 days old [16, 17]. The reduced capacity of the liver and kidney to detoxify and eliminate chloramphenicol explains the development of toxic effects in neonates, especially in premature neonates given body-weight scaled chloramphenicol doses [18]. Another possible reason is that children may show development-associated toxicity that adults do

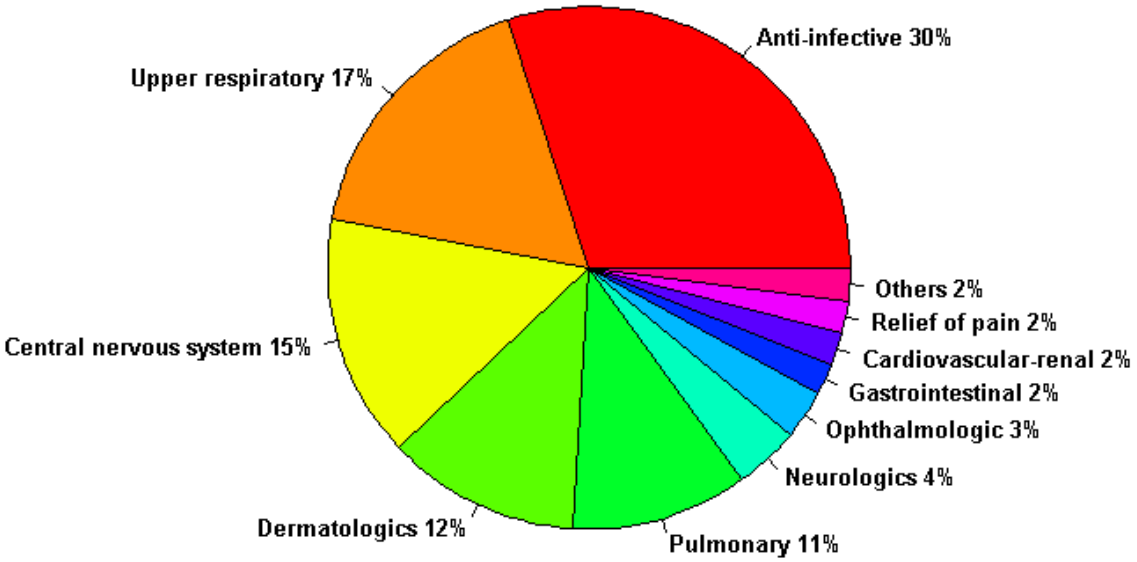


Figure 1-1. Percentage of off-label drug use in major drug category in pediatric population

Figure shows most often prescribed off-label drug categories in pediatrics.

Source: Figure based on data from Bazzano, A.T., et al., *Off-label prescribing to children in the United States outpatient setting.* Acad Pediatr, 2009. **9**(2): p. 81-8.

not experience. For example, tetracycline-induced discolorations of teeth only affect developing enamel during the period of calcification [19]. Therefore tetracycline is today no longer prescribed for young children and pregnant women. These two examples illustrate pharmacokinetic (PK) and pharmacodynamic (PD) differences between adults and young children resulting in very disparate responses to drug therapy. However, the differences in how PK and/ or PD interact in a fast growing child have not been thoroughly investigated. In such cases, linear predictions from adult doses (e.g., dose extrapolation on the basis of linear scaling per body surface area [BSA] or per body weight [WT]) are not always suitable solutions.

The American Academy of Pediatrics Committee on Drugs has stated: "There is a moral imperative to formally study drugs in children so that they can enjoy equal access to existing as well as new therapeutic agents" [20]. To correct the situation of off-label dosing, a series of laws and regulations have been enacted regarding the integrity of the research and development of therapies for pediatric patients. Recent legislative incentives and regulations for pediatric drug development are summarized in **Table 1-1**.

The Pediatric Labeling Rule issued by the FDA in 1994 encourages sponsors to review existing data and provide the appropriate labeling information for pediatric use if the course of the disease and the expected drug responses are similar between adult and pediatric populations. However, only a small increase in the number of applications with supplemental pediatric labeling information was noticed following the Pediatric Labeling Rule [21]. The FDA Modernization Act (FDAMA) issued in 1997 included financial incentives for pharmaceutical companies to conduct pediatric studies. This act offered an additional 6 months of marketing exclusivity for a drug tested in pediatric studies and also led to the development of an annual Pediatric Priority List by the FDA, which consists of approved drugs that need new pediatric use information. In 1998, the FDA published the Pediatric Rule, which required that any new drug application contain the data from pediatric testing unless this drug was not going to be used in a substantial number of pediatric patients. Under the Pediatric Rule, the FDA also required drug companies to perform pediatric studies for a marketed drug when the drug was used or intended to be used in pediatric patients but lacked labeling information where the drug might cause significant risks. The change of an active ingredient, formulation, dosage or route of administration of a drug may trigger this rule [21]. However, the Pediatric Rule

Table 1-1. Summary of regulatory incentives and regulations for drug development in pediatric populations

Year	Regulatory incentives or regulations
1994	Pediatric Labeling Rule
1997	FDA Modernization Act (FDAMA)
1998	Pediatric Rule
2002	Best Pharmaceuticals for Children Act (BPCA)
2003	Pediatric Research Equity Act (PREA)
2007	FDA Amendments Act

was overturned by a federal court in 2002. More recently, the Best Pharmaceuticals for Children Act (BPCA) took effect in 2002 and aimed to improve pediatric drug development as well as to encourage testing of new drugs and off-patent drugs for use in pediatric populations. The Pediatric Research Equity Act (PREA) was signed into law in 2003; it authorizes the FDA to require clinical studies in pediatric populations for new drugs and biological products targeting pediatric patients. To reauthorize the BPCA and PREA, the FDA Amendments Act was enacted in 2007; it extended the 6-month additional market exclusivity for patented drugs when clinical studies are being conducted in pediatric populations.

Role of Pharmacokinetics and Pharmacodynamics in Pediatric Optimal Dosing

In the previous section, a couple of examples explicating different drug responses derived from large variability in pharmacokinetics and pharmacodynamics between adults and children were mentioned. These differences also exist within the gestational age variance of premature infants. Premature infants are not a homogenous group; their maturational process after birth may follow various patterns. Developmental changes in body size and function are rapid and continuous. A variety of changes in physiological and biochemical processes resulting in age-dependent differences in drug disposition (i.e., absorption, distribution, metabolism and excretion) and drug response also exist in these infants.

PKPD and optimal dose selection

Pharmacokinetics describes the effect of the body on drugs, including the process of absorption, distribution, metabolism and excretion (ADME) of a drug over a certain period of time. Pharmacodynamics describes the relationship between drug concentrations and the magnitude of drug effect at the assumed site of drug action. In other words, PK and PD may be simply defined as ‘what the body does to the drug’ and ‘what the drug does to the body’ [22]. The PK and PD characteristics may explain the clinical responses of certain dosing regimens of a drug independently or their combined effect—i.e., dose–exposure–response relationship can play an important role in deciding the optimal dosing regimen with maximal efficacy and minimum undesirable drug effects. The exposure-response relationship for pediatric clinical pharmacology may not be the same as for adults. The understanding of the PK and/or PD relationship can provide a rational and scientifically based framework for the determination of the optimal dosing regimen in pediatric pharmacotherapy. When a similar concentration-response relationship can be assumed between pediatrics and adults and when PK is the major factor contributing to differences in drug responses, optimal dose selection can be determined based on the PK characteristics of the drug. One example is topiramate, which is a recently approved drug for the monotherapy of seizures in pediatric patients 2-10 years old [23]. Topiramate has been previously approved as monotherapy in patients > 10 years of age or as adjunctive therapy for pediatric patients 2-10 years of age. In this study, a similar exposure-response relationship was proven between pediatric patients

and adults through PK/PD modeling and simulation. Therefore, the dosing regimen for pediatric patients 2-10 years old was derived based on PopPK modeling and simulation without additional clinical trials. Meanwhile, dose optimization also may be determined by the drug's PK and PD properties simultaneously when both PK and PD characteristics are changing. For example, dosing recommendations for sotalol have been derived when both PK and PD data were considered [24]. In this study, different concentration-QT prolongation relationships were revealed through population PK/PD analysis between neonates and older patient groups. Based on the findings, both solatol's systemic exposures and PD responses were included in the development of sotalol dosage recommendations in different pediatric age groups to achieve maximum efficacy and minimum safety concerns.

Age-dependent changes in PKPD

Absorption

Absorption of oral drugs can be affected by gastric acidity, gastric motility and emptying time, as well as the length of the gastrointestinal tract. Gastric acidity, which is decreased in neonates and infants, can enhance the absorption of acid labile drugs, such as benzylpenicillin and ampicillin, and decrease the absorption of acidic drugs, such as phenobarbital and phenytoin [25]. Gastric pH then changes gradually during maturation and slowly reaches adult levels approximately 2 years after birth. In addition, gastric emptying is much slower in premature infants compared to term infants and older children [26]. Thus bioavailability of oral drugs in premature infants can be very different from that in term infants and older children.

Distribution

Drug distribution is also different in neonates and infants and is greatly influenced by total body water and protein binding. Total body water is approximately 85% of the body weight in premature neonates and 70-75% in term neonates compared with 50-55% in adults. As a result, a larger dose (per body weight) is required for polar compounds to achieve therapeutic concentrations in neonates [25]. Additionally, volume of distribution can be affected by protein binding. Due to the lower levels, protein binding of drugs is generally reduced in neonates, especially premature neonates, as compared to that in older children and adults. Reduced protein binding also leads to a higher fraction of unbound plasma drug concentrations, which results in a larger volume of distribution for medications used in newborns. For example, it has been reported that with administration of theophylline a lower protein binding and a larger volume of distribution were observed in preterm neonates than in older children and adults [27].

Metabolism

Drug metabolism defines the biotransformation of endogenous and exogenous compounds in the body. It may occur in the liver, kidneys, intestines, lungs and blood

cells, but hepatic metabolism is the primary pathway for most drugs' metabolism. Hepatic metabolism is usually catalyzed in the liver by the cytochrome P450 (CYP) enzymes (Phase I reactions), and/or the Phase II enzymes, such as uridine diphosphate glucuronosyl transferase, sulfotransferase, methyltransferase, glutathione S-transferase and N-acetyltransferases. Metabolism generally enhances drug excretion by transforming xenobiotics and drugs into a more water-soluble form [28]. These metabolic pathways are generally underdeveloped at birth and change dramatically from birth to adulthood. The rates of maturation for metabolizing enzymes are much slower in premature neonates than in term neonates [25].

Although liver expression of major Phase I enzymes, such as the CYP1A, CYP2C, CYP2D, CYP2E and CYP3A subfamilies, are generally very low in neonates and reach adult levels within 2 years after birth, the age-dependent developmental trajectory might be different between different type of enzymes. For example, CYP3A4 is the most abundant isozyme in the CYP family in the adult liver and accounts for the metabolism of more than 30% of all drugs, while CYP3A7 is the predominant enzyme of the CYP3A family in the fetal liver [29, 30]. CYP3A7 activity remains maximal within one week after birth but then decreases to a very low level. Concurrently, CYP3A4 enzyme activity increases after birth, reaching approximately 30-40% of adult levels by one month of age, and actually exceeds adult levels by two years of age [29, 31]. The CYP2C enzymes are an important subfamily of CYPs and are involved in the metabolism of ~20% of clinically used drugs in adults [32]. Of them, CYP2C9 and CYP2C19 are two predominant isozymes. CYP2C9 protein and activity levels are comparable to adult values at birth, while CYP2C19 enzyme activity is only 12 to 15% of adult values at birth and reaches adult levels after 10 years of age [33]. Very low levels of CYP2D6 protein can be detected in the fetal liver and are associated with the O-demethylation activity of dextromethorphan [34, 35]. CYP2D6 activity increases rapidly after birth and reaches approximately 30% of adult levels by the first month of life [34]. Adult levels of CYP2D6 enzyme activity may be reached by one year of age [36]. CYP1A2 enzyme activity is negligible at birth and remains very low until one to three months after birth. Its activity increases to approximately 50% of adult values by one year of age [37].

Phase II enzymes also contribute significantly to the elimination of many clinically used drugs. Though the information for the ontogeny of Phase II metabolizing enzymes is still limited, current literature indicates that developmental changes of many Phase II enzymes occur throughout infancy and usually exhibit isoform-specific maturation process [38]. Therefore, reduced ability to eliminate exogenous and endogenous compounds also may be observed in neonates and infants because of the low conjugation capacity of the Phase II enzymes.

As a result, drugs that are metabolized by the Phase I and Phase II enzymes may exhibit a low systemic clearance and prolonged elimination half-life in neonates; but a higher weight-corrected dose may also be needed during infant or children period when metabolizing enzyme activities reach adult levels due to nonlinear scaling requirements between children and adults. Although expression of some hepatic enzymes might appear to be associated with birth, both postconceptional and postnatal development can affect hepatic drug clearance in neonates. Postconceptional age is considered a more

physiologically appropriate factor to explain maturation-dependent drug metabolism, as it explains the maturation and developmental process of drug metabolic pathways both prior to and after birth [39].

Xenobiotics and their metabolites can be excreted from cells by Phase III transporters, most of which belong to the ATP-binding cassette family (ABC) or the solute carrier transporter family. Phase III transporters are present in the cell membrane of many tissues, such as liver, intestine, kidney and brain, and can provide a barrier against xenobiotic entry or move various endogenous and exogenous compounds in and out of cells. For example, Phase III transporters are involved in the excretion of drugs and their metabolites into the bile for hepatic elimination and are essential components contributing to the overall hepatic clearance.

The ABC transporters are a superfamily of large membrane proteins that can actively transport a variety of compounds through membranes in an ATP-dependant manner [40], and human ABC transporters can be divided into seven subfamilies. Among these subfamilies, multidrug resistance P-glycoprotein (P-gp, coded by the MDR1/ABCB1 gene) and multidrug resistance protein 2 (MRP2, coded by ABCC2 gene) appear to be the most relevant transporters for the hepatobiliary elimination of xenobiotics in humans. P-gp exists in many major organs including liver, kidney, small intestine, and blood brain barrier and plays an important role in the disposition of a variety of hydrophobic and cationic drugs. MRP2 is highly expressed in the liver, intestine, and kidney and transports a range of drugs conjugated to glutathione, sulfate, or glucuronate into bile [40].

Although Phase III transporters may significantly influence drug absorption and elimination, the developmental changes of drug transporters remains largely unknown. For example, loperamide (a P-gp substrate)-induced respiratory depression due to central nervous toxicity has been reported in infants despite its apparent safety in adults [41, 42]. However, it is unclear whether or not this toxicity is caused by the lower expression of P-gp at the blood brain barrier in infants. Additionally, one recent study has reported that P-gp is expressed in a developmental and cell-specific manner in the human central nervous system [43]. Thus, it is possible that the expression pattern of P-gp at the blood brain barrier might affect the uptake or excretion of drugs in the central nervous system in neonates. Similarly, potential MRP2-related drug toxicity has been reported in pediatric populations [44-47]. Ceftriaxone, an antibiotic often used to treat lower respiratory infections or acute otitis media in the pediatric population, is partially eliminated unchanged into the bile through the MRP2 transporter. An increased risk of cholestasis has been reported in children receiving the administration of ceftriaxone, which may be caused by the lower expression of MRP2 in pediatric patients, thus leading to an accumulation of ceftriaxone in the hepatocytes. However, further investigations are still needed to confirm the underlying mechanism of this adverse drug reaction.

Excretion

Renal elimination mechanisms include glomerular filtration, active secretion and tubular reabsorption. The maturation of glomerular filtration differs between term and preterm infants and may affect the drug's renal elimination. Glomerular filtration rates are only approximately 30% of adult function at birth and mature during infancy with an exponential function that asymptotically reaches adult values [39]. Adult levels are reached within 1 year of age for a full term infant [48]. Premature infants have a much smaller increase in glomerular filtration rate during the first 3 days after birth than term infants [49]. Various transporters such as ABC transporters, organic anion/cation transporters (OATs/OCTs) and the peptide co-transporter PEPT2, are involved in renal secretion and reabsorption [50]. However, little information is known about the maturation process of these transporters in the kidney.

Developmental changes in pharmacodynamics

Both PK and PD processes contribute to the difference in the safety and efficacy of a drug between pediatric populations and adults. A decision tree for conducting pediatric studies was proposed by the FDA in 2003 [51]. The FDA guidance for bridging efficacy studies recommends the evaluation of disease progression, drug response and concentration-response relationship between adults and pediatric populations. Although currently there is little known regarding the developmental changes in PD responses, a number of studies have reported age-dependant changes in the concentration-response relationships of some drugs. For example, Tran et al. reported that the antisecretory effect following a single oral lansoprazole administration was increased in infants younger than 6 months compared to that in older children and adults [52]. Similarly, age-related differences in PD responses have been reported for tacrolimus [53], cyclosporine [54], sotalol [24], rocuronium [55], and warfarin [56].

These maturation processes in premature neonates and infants have significant impact on a drug's dose-concentration relationship, which may lead to profound age-dependent differences in PK and PD responses. Drugs, including newly developed and established ones, must be investigated in children to determine their safety and efficacy in corresponding age groups. However, only therapeutic studies or routine care studies are considered ethically acceptable research in children. New research approaches have been proposed to overcome those challenges.

Challenges of Clinical Studies in Pediatric Patients

As mentioned earlier, the prevalence of off-label medications use in neonates and young infants is most likely due to the greater challenges in conducting clinical studies in this population as well as inadequate knowledge of developmental pharmacology [3, 4, 57]. This situation is even worse for premature infants. Some unique challenges in clinical studies in this population are illustrated in **Figure 1-2**. In the United States the preterm birth rate has been on the rise for more than two decades, with a rate of 12.3% in

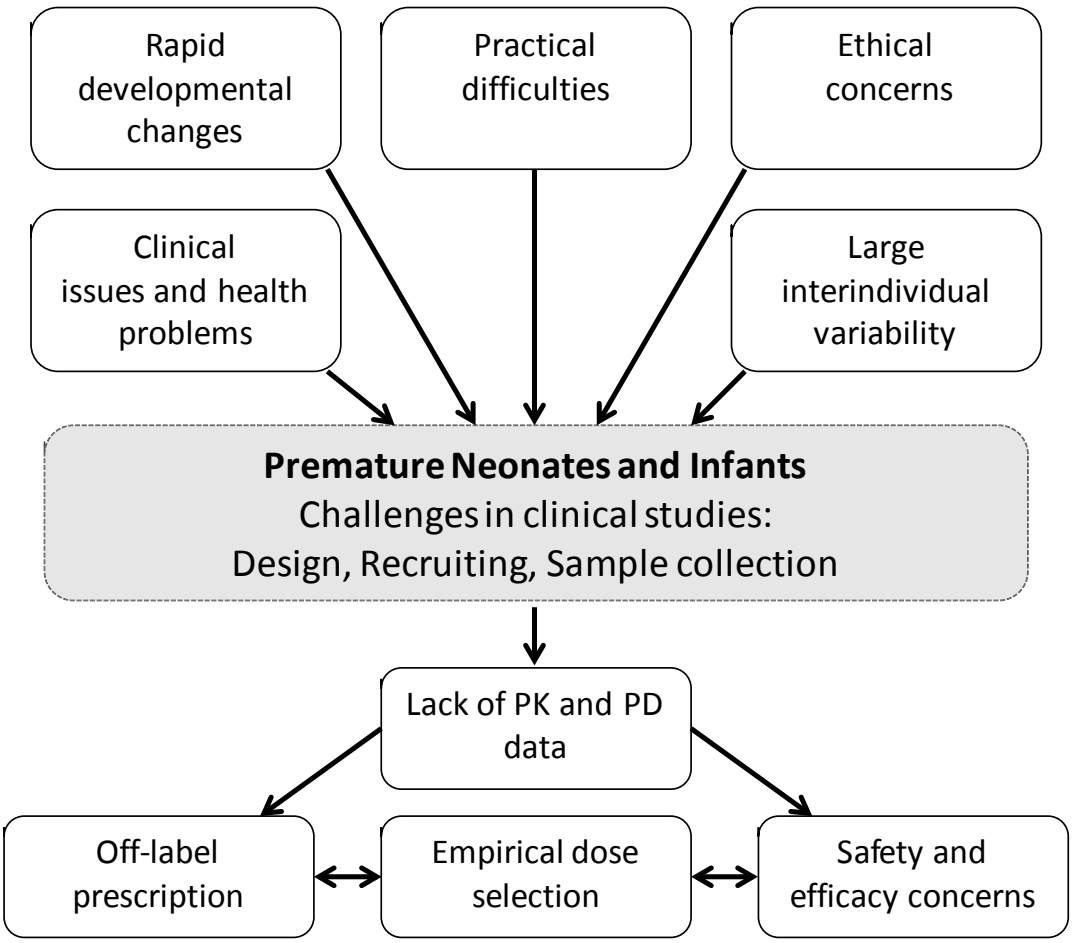


Figure 1-2. Illustration of unique challenges of clinical studies in premature neonates and infants

PK = pharmacokinetics
PD = pharmacodynamics

2008 [58]. Despite advances in neonatal care, premature neonates and infants remain a major group for morbidity and mortality in all infants. As this population expands, researchers are gaining more and more awareness for the need for rational and scientific-based pharmacotherapy other than empirical-based dosing.

Health problems in premature infants

Compared to term infants, premature infants face much higher rates of health issues after birth. Briefly, common problems include but are not limited to: 1. low birth weight, 2. feeding problems, 3. severe infections, and 4. immaturity of major organs or organ systems. For example, almost all newborn babies with body weight < 1000 g or with gestational age < 29 weeks have breathing problems to some extent due to underdeveloped lungs and a resulting insufficient oxygen supply. Special care is often needed to support their lives for weeks to months. Medical and ethical concerns highly restrict the performance of clinical research in this patient group. As a result, new drug testing usually excludes pediatric populations from therapeutic research for the establishment of evidence-based medicine, except for therapeutic studies or routine care studies.

Limited sample volumes in pediatric studies

The volume and frequency of blood sample collection in infants and small children are usually very limited. While no consensus has been reached on the appropriate level of pediatric blood collection, various criteria have been proposed to minimize the volume of blood withdrawn in pediatric studies. For example, the Partners Human Research Committee (PHRC) recommended that no more than 3 mL/kg of blood may be drawn per 8-week period for research purposes in children [59]. The European Commission recommended that no more than 3% of the total blood volume may be drawn for research purposes over a 4-week period in neonates and children [60]. Therefore, clinicians and researchers are urged to pay special attention on developing sensitive assays that allow for the determination of drug concentrations by using a small volume of blood samples.

In addition to the issues addressed above, there are numerous barriers to conducting clinical studies in premature infants, including scientific, ethical, practical and financial limitations. All of these challenges hinder further evaluation of new drugs and established treatments in children. Thus, research to gain a broad understanding of developmental changes in pharmacokinetics and pharmacodynamics in young children needs to take advantage of appropriate research tools to develop the best dosing strategies for young children.

Some Research Strategies in Pediatric Studies

Quantitative assays: liquid chromatography tandem mass spectrometry (LC-MS/MS) technique

In order to overcome some of the limitations of PK studies in young children, innovative analytical approaches to facilitate pharmacokinetic evaluation in children are needed. Liquid chromatography coupled with tandem mass spectrometry detection (LC-MS/MS) has developed into a powerful, analytical approach because of its robustness, high sensitivity, selectivity and accuracy, allowing for reliable drug and metabolite quantification at low therapeutic levels. Given the fact that 90% of medications used in the NICU are off-labeled, there is a great need for the study of medication commonly used by newborn infants to ensure their safety and efficacy. The ability to determine multiple analytes within one small volume of biological fluids (mostly plasma) at one time is highly favorable in pharmacokinetic studies of newborns. This feature is of great benefit because it enables researchers to gain the maximum amount of PK information in different drugs through the quantitation of multiple drugs from a single micro volume of plasma, thus minimizing the burden of invasive venous punctures in infants involved in pharmacokinetic studies. A number of quantitative assays using LC-MS/MS methodology have been previously described [61-67]. The focus on this analytical technology allows us to rapidly establish flexible assays to quantify drugs of interest classified by their physicochemical properties. The most commonly prescribed medications for premature neonates in the NICU and for infants were placed on the top of the priority list to be studied. Among them, acetaminophen, caffeine, phenytoin, ranitidine and theophylline were included in our current study.

Population pharmacokinetics

Modeling and simulation (M&S) are methods that have been widely used in drug development to support study design, data analysis, and study decision making. Modeling allows for a quantitative description of pharmacokinetic and/or pharmacodynamic properties of the investigated drugs by analyzing the observed data from preclinical and clinical studies. It is usually expressed in a simplified manner, such as focusing only on the important factors, to characterize a system or process [68]. Simulation refers to the use of established pharmacokinetic and/or pharmacodynamic models to predict future outcomes that have not been investigated experimentally. Model-based drug development has been widely recognized as an invaluable tool in clinical research and is also recommended by the FDA to be applied to drug development to improve knowledge management and decision making [69].

Traditional PK and PopPK

The traditional PK approach is widely used in the early phase of drug development. It usually involves a small group of subjects who are followed with an

intensive pre-defined sampling scheme. Some special populations, such as the elderly, pediatric patients or severe anemia patients, are not suitable for these study interventions due to ethical and medical concerns. To characterize the pharmacokinetics of the studied population in a traditional PK study, a standard two-stage (STS) approach is applied: First each individual's PK parameters is estimated from dense sampling, and then population parameters are expressed as mean and standard deviation of the estimated individuals. For the most part, mean values of the parameters have little or no bias, but estimation of between-subject variability is biased due to the small number of studied subjects. A rich data set will provide the most accurate information.

Application of population pharmacokinetics (PopPK) was first introduced in the 1970s by Sheiner and colleagues [70]. Today, PopPK has become a standard approach to investigate PK data during drug development and evaluation. PopPK is defined as the study of the sources and correlates of variability in drug concentrations among individuals who represent the target patient population receiving the clinically relevant doses of a drug [71]. Compared to traditional PK, it has several advantages in data analysis. It allows for simultaneously analyzing pooled data from multiple subjects with a relatively flexible sampling scheme. The number of samples from each subject can be sparse, dense or a combination of sparse and dense data (mixed sampling), and sampling times do not have to be same from all patients. This can greatly reduce the individual burden of blood sampling as well as adapt to the patients' convenience for blood sampling. It allows simultaneous estimation of the typical values of PK parameters in the targeted population and identifies the sources and magnitude of the variability as well. Through the population analysis, variability in the PK and PD response due to predictive factors or covariates (such as demographics, genetics, concomitant drug administration, pathophysiological conditions, disease status, food effects, formulation, etc.) can be considered and identified. The analysis also provides empiric Bayesian estimates of individual parameters and parameter associated between-subject variability, which subsequently can be used for individualized dosing strategy.

PopPK and nonlinear mixed effects model

Nonlinear mixed effects modeling is often used to perform PopPK analyses. The typical model includes fixed effects (θ) and two levels of unexplained variability (random effects), between-subject variability (η) and residual variability (ε). η describes the differences between the typical population parameter θ and the individual-specific parameter estimate. Its distribution is assumed to be normal, with mean of 0 and variance of ω^2. ε describes the difference between model predicted concentration and observed concentration, with expected distribution of mean 0 and variance σ^2 [72]. Residual error could come from incorrect dosing, sampling record errors, assay error, within-subject variability, or model misspecification.

The PopPK model process includes two primary components: a structural model and a covariate model. The structural model, which most times is also referred to as the base model, is the simplest model ready for stepwise model building. Generally, the structural model is the model that can best describe the data without any covariate.

However, in some special cases, it's widely believed that a relationship between model parameters and a predictive factor can also become part of the structural model *a priori*. An example is size adjustment using the fixed-exponential allometric model *a priori* in pediatric PopPK analysis [73]. The covariate model building process explores and identifies factors that are important in explaining/reducing between-subject variability and residual variability. In premature neonates and infants, age (gestational age, postnatal age [PNA], postconceptional age [PCA]) and size (body weight, birth weight [BW], body surface area, lean body weight) are two classes of significant factors correlated to maturation process. Other examples of covariates are concomitant medication, genotype, and biomarkers. Covariates can be added into the base model in a stepwise fashion, either with proportional, exponential, fractional or additive relationships, depending on their nature (categorical or continuous) and their correlation to parameters [74, 75].

Several software packages are capable of handling population-based PKPD analysis [76], including NLME (Bell Labs, Murray Hill, New Jersey), WinNONMIX® (Pharsight Corporation, Mountain View, California), Monolix (LIXOFT, Orsay, France), WinBUGS (Bayesian via Markov Chain Monte Carlo [MCMC] methods, MRC Biostatistics Unit, Cambridge, UK), SAS (SAS Institute, Cary, North Carolina), S-ADAPT (Biomedical Simulations Resource, Los Angeles, California), and S-Plus (TIBCO Software, Palo Alto, California). The most widely used software today for the characterization of PopPK, however, is NONMEM® (distributed by ICON Development Solutions, Ellicott City, Maryland).

Application of modeling and simulation in pediatrics

The modeling and simulation (M&S) approach has been widely used in all phases of drug development and drug evaluation, from preclinical to post-marketing (**Figure 1-3)** [77]. Today it has become an integral part of the drug development process and regulatory decision making [78]. In a review of 198 submissions between 2000 and 2008 to the US FDA [79], the number of cases involving M&S analyses (pharmacometric approach) had increased 6-fold over 9 years. Among them, 26% of the submissions included pediatric studies.

The core component of pediatric studies is to provide optimized dosing regimens for safety and efficacy with both new and established medications. On the one hand, scientific-based drug development and pharmacotherapy are greatly needed and have been mandatory by legislation [80]; on the other hand, pediatric studies are hampered to a great extent by ethical, medical and practical issues. The advantages of M&S are to give added value to pediatric drug studies for its ability to overcoming such challenges in this population group. Due to recent regulatory incentives for the conduct of pediatric studies, population-based M&S plays an increasingly important role in pediatric drug development [81]. In the current work, we focus on the effect of population-based M&S on data analyses, sample size estimation and selection of dosing regimens.

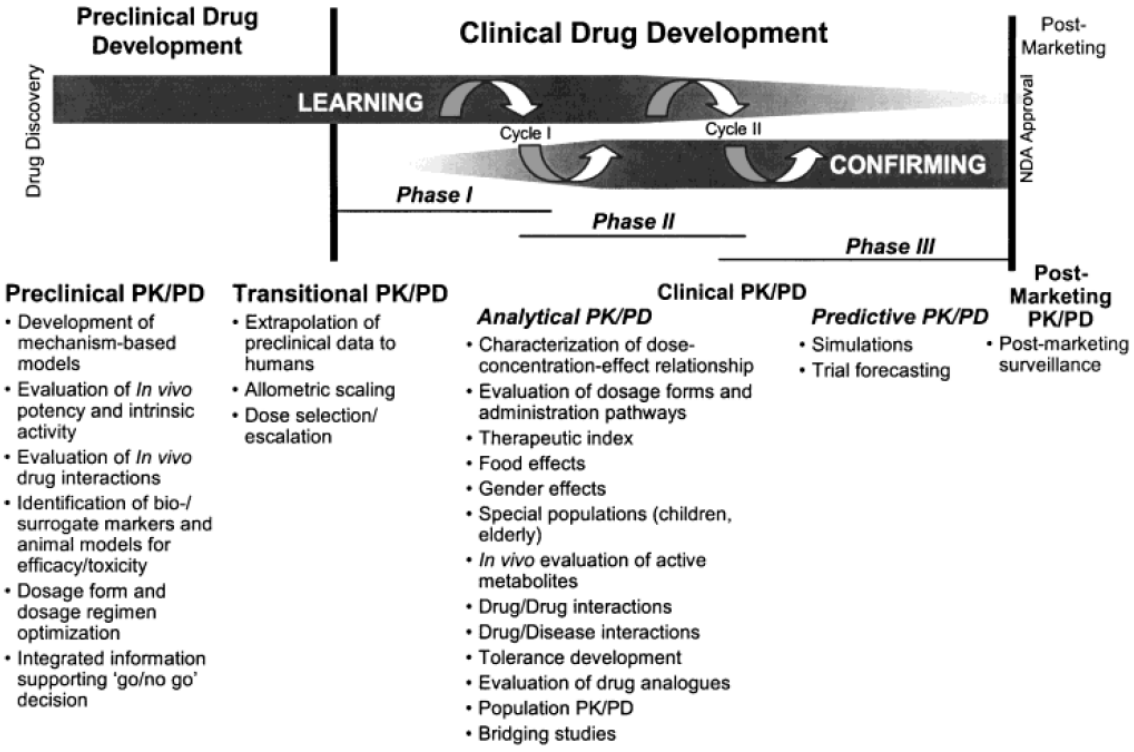


Figure 1-3. Potential applications of modeling and simulation concepts during preclinical and clinical drug product development

Source: Reprinted with permission. Meibohm, B. and H. Derendorf, *Pharmacokinetic/-pharmacodynamic studies in drug product development.* J Pharm Sci, 2002. **91**(1): p. 18-31.

Population-based modeling: sparse and unbalanced sampling

As we discussed, one of the specific advantages of population-based M&S is the capability of handling sparse and unbalanced data from a large, heterogeneous group of patients, which is the situation in neonates and infants studies [73]. Because of ethical and medical concerns, young children are well protected by minimizing the invasive samplings and all blood specimens are usually only collected for a therapeutic or diagnostic purpose. Therefore, the number of blood samples from each individual is very limited and usually doesn't follow a fixed sampling schedule. The successful application of M&S makes the study of this critically ill population (NICU patients) more feasible. A number of PopPK studies were conducted in children using sparse data with unbalanced design, or they took advantage of therapeutic drug monitoring (TDM) data, and the developed models and results were successfully translated into clinical therapeutic decision making, thereby benefitting patients with optimal dosing regimens. Some of the examples include vancomycin [82-84], phenytoin [85], midazolam [86], aminoglycosides (gentamicin, tobramycin, netilicin, amikacin) [87] and sotalol [24]. Identification of significant factors contributing to the variability of parameters can be especially important in premature infants where rapid developmental changes occur over relatively short periods of time, resulting in a large variability in drug disposition. The developed models can be used in the future for dosing regimen optimization by relating PK parameters (such as CL) and patient demographic factors (such as body weight and, age) to criteria for therapeutic safety and efficacy.

Clinical trial simulation

Clinical study designs can be explored by simulation based on PKPD modeling. M&S as a powerful tool for rationale decision making provides the capability for careful design and planning in pediatric studies. Clinical trial simulations closely depend on population pharmacokinetics. They allow researchers and clinicians to explore situations that have not been investigated before—thereby gaining insight into a "new world," for example, extrapolating results from animals to humans. Another advantage of clinical trial simulations is to help researchers tailor design factors by comparing and investigating simulation results as if in the "real world"—including, for example, optimal sampling design for a trial, optimized sample size estimation, or optimized dosing regimens. M&S provides a scientific framework for efficient decision making, thereby increasing the probability of success in clinical studies while minimizing risk and cost.

Optimal study design

A carefully designed clinical study will improve the probability of "success" by comparing and assessing the impact of different design factors that may affect the outcome, such as dosing regimen, sample size, number of drop-outs and trial duration, with considering uncertainties [88]. Sample size estimation is a key factor for a successful study with adequate power and reliable results but involving a minimum number of patients (sample size) in order to minimize the trial duration, cost, and the

potential risk imposed to the patients. A minimum but adequate sample size is one of the most important aspects of optimal study design in pediatric patients due to the ethical limitations in this population. However, sample size estimation has not been included in most population pharmacokinetic studies. Through the M&S based upon available PKPD data and/or prior knowledge, the required number of subjects can be estimated to best reflect the study objectives and characterize pharmacokinetics in specific populations. We will illustrate this application in Chapter 3.

Rapid developmental change is another feature associated with pediatric patients, especially with neonates and infants, which leads to large between-subject and within-subject variability in premature neonates and infants. If developmental changes are not included during the study design, the clinical trials are likely to fail due to the lack of knowledge of drug action during the development process [68, 89]. Simulations coupled with PopPK models allow testing the different designs and the impact of uncertainty on the outcome of the study in a computer-based, virtual environment; they provide researchers chance to gain insight into the results before the study is actually performed. Mouksassi et al. [90] used PopPK and clinical trial simulations to select dosing regimens for a phase I study of teduglutide in pediatric patients with short-bowel syndrome. In their study, realistic covariate input specific to the targeted patient population was simulated and used to evaluate dosing strategies under various age-weight, pathophysiological conditions thereby determining safety and efficacy in this patient population. Thus in the pediatric group, M&S is a useful tool to optimize the study design by incorporating the growth effect and maturational changes, and maximize the likelihood of achieving target exposure in the real clinical setting. Meanwhile, appropriate sample size will ensure a successful study with fewer patients exposed to the investigational procedures, which is also meaningful for the ethical and practical considerations when conducting a pediatric clinical trial.

Optimal dosing regimen

Traditionally, approval of a new drug application by FDA was primarily determined by reviewing the medical and statistical data. As described previously, an important advance in clinical drug investigations is the incorporation of population-based M&S into an approval decision [78].

Among all types of decisions making, the majority of cases were relevant to dosing selection based on quantitative benefit-risk assessments. Today, pharmacometrics allows for dosing regimens to be based on modeling and simulation analyses before they are thoroughly studied in phase III clinical trials, or they are supported by pharmacometric analyses as confirmatory evidence for supporting labeling information.

Modeling and simulation can be useful in establishing optimal dosing strategies and increasing the successful probability of a clinical study. Empirical Bayes estimation of individual pharmacokinetic parameters acquired by modeling analysis, combined with individualized measurement, has been successfully applied in pediatric PK studies for individual optimal dosing selection. Simulation of PKPD can be performed to predict the

drug concentrations or responses under a "real world" condition with different dosing regimens. Thus appropriate study design and dosing strategy can be proposed for pediatric studies based not only on empirical assumptions but also on a model-based approach, **Figure 1-4 s**hows a schematic illustration of this approach. Such approaches have been evaluated in multiple studies for regulatory decision making [91-93]. Clinical trial simulation allows the utilization of population PK/PD models along with the integration of study design, patient demographics and disease status. As a result, optimal design can be selected and dosing strategy may be evaluated in various conditions. For example, the approval of levofloxacin dosage to treat anthrax in children was based on pharmacometric analyses with M&S since no clinical trials could be conducted with the recommended dosing regimen [94].

Summary and Specific Aims

Research efforts focused on optimized dosing strategies for safe and effective use of medications in premature infants is needed to improve our understanding of drug disposition in this population group. The utilization of advanced analytical assays will allow for pharmacokinetic studies of drugs commonly used in premature infants. The aims of this present work were to investigate research strategies in the pharmacokinetic study of drugs used in premature neonates and infants, including bioanalytic assay development, trial design factor investigation, PopPK model development and dosing regimen assessment. These processes were aimed at developing optimized dosing regimens for premature neonates and infants.

In specific aim 1 (discussed in Chapter 2), to enhance our knowledge on pharmacokinetics of commonly used drugs in premature neonates, we developed and validated an LC-MS/MS method for the simultaneous determination of commonly used medications in the NICU, including acetaminophen, caffeine, phenytoin, ranitidine, and theophylline, in small volume human plasma specimens of 50 μL [95].

In specific aim 2 (discussed in Chapter 3), we explored sample size requirements for observational PopPK studies in premature neonates and infants using theophylline as a model drug. A full model-based simulation approach was applied with prior information and between-subject variability and residual variability. We evaluated the accuracy, precision and power of parameter estimation and also investigated the effect of sample size on the detection of significant covariates.

For specific aim 3 (discussed in Chapter 4), we developed a PopPK model of caffeine in premature neonates and evaluated the change of PK parameters throughout infancy. The developed model was subsequently used for a dose-optimization study by simulation, particularly to simulate the distributions of steady state concentrations at different dosing regimens for various age/body size groups, which provided the rationale for age/weight specific, optimized dosing regimens.

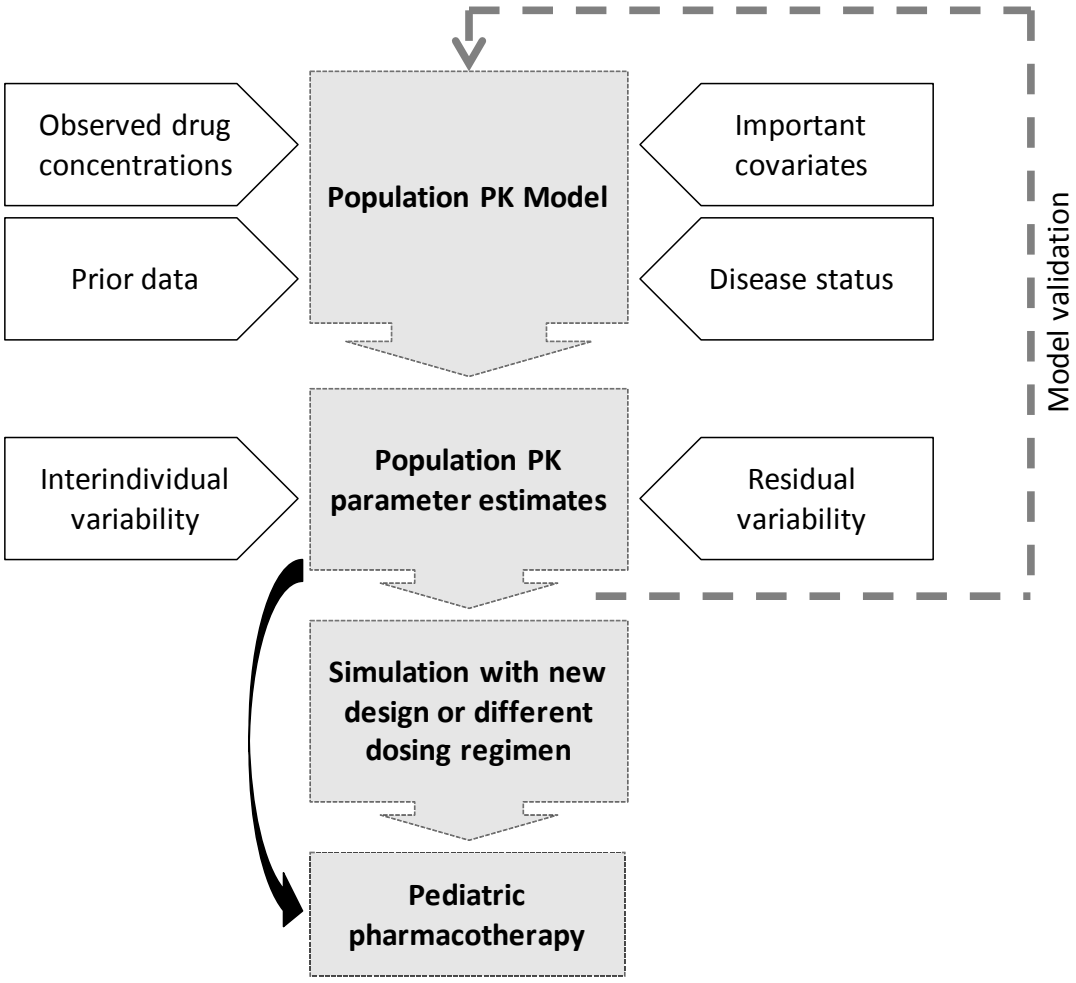


Figure 1-4. Application of PopPK modeling and simulation in pediatric pharmacotherapy

PK = pharmacokinetics.

CHAPTER 2. A TANDEM MASS SPECTROMETRY ASSAY FOR THE SIMULTANEOUS DETERMINATION OF ACETAMINOPHEN, CAFFEINE, PHENYTOIN, RANITIDINE, AND THEOPHYLLINE IN SMALL VOLUME PLASMA SPECIMENS*

Introduction

Premature infants (gestational age less than 37 weeks) are considered a vulnerable patient population due to their immaturity at birth. Born at different gestational ages, they experience rapid growth and continuous developmental changes in body size and composition as well as organ size and function. Different stages of maturation and different maturational trajectories for the physiological and biochemical processes that govern drug disposition (i.e., absorption, distribution, metabolism, and excretion) result in tremendous inter-individual pharmacokinetic variability, leading to very disparate responses to drug therapy [96, 97].

Acetaminophen, caffeine, phenytoin, ranitidine, and theophylline are widely used in the pharmacotherapy of premature and term neonates. Acetaminophen, or paracetamol, is an effective and widely used analgesic and antipyretic medication in infants [98]. Caffeine and theophylline are both used in the treatment of neonatal apnea in premature infants [99-101]. Ranitidine is frequently used for the reduction of intragastric acidity in conditions such as pathological gastro-oesophageal reflux or stress ulcer prophylaxis in critically ill infants, the latter being a common side effect of steroid treatment in premature infants with bronchopulmonary dysplasia [102]. Phenytoin is applied as a second line medication for the pharmacotherapy of seizures in patients with treatment failure on phenobarbital therapy [103]. Only limited information is currently available on the pharmacokinetics of these medications in premature neonates. Ethical and practical constraints in sample collection from this patient population limit the number and volume of blood specimens available for pharmacokinetic evaluations [104].

In recent years, high performance liquid chromatography with mass spectrometry detection (LC-MS/MS) has become the standard analytical methodology in pharmacokinetic evaluations due to its robustness and high sensitivity. LC-MS/MS allows for reliable drug and metabolite quantification even within the confines of small sample volumes in pediatric studies [104]. A number of quantitative assays using LC-MS/MS for the above mentioned drugs have been previously described. These methods, however, are limited to the quantification of one specific drug per assay, and many do not have sufficient sensitivity to quantify therapeutic drug concentrations in small volume plasma specimens [61-67]. Due to the limitations in sample volume, only

*Adapted with permission. Zhang, Y., et al., *A tandem mass spectrometry assay for the simultaneous determination of acetaminophen, caffeine, phenytoin, ranitidine, and theophylline in small volume pediatric plasma specimens.* Clin Chim Acta, 2008. **398**(1-2): p. 105-12.

an assay that can simultaneously determine multiple drugs concurrently used in the pharmacotherapy of premature neonates was deemed feasible to support pharmacokinetic studies in this population. Thus, in the present study, we developed and validated an LC-MS/MS method for the simultaneous determination of acetaminophen, caffeine, phenytoin, ranitidine, and theophylline in small volume human plasma specimens of 50 µL.

Materials and Methods

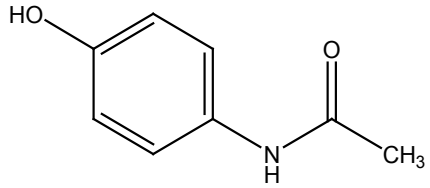
Chemicals and reagents

All chemicals used including acetaminophen ($C_8H_9NO_2$, 99.0%, MW 151.2), caffeine ($C_8H_{10}N_4O_2$, 99.9%, MW 194.2), phenytoin ($C_{15}H_{12}N_2O_2$, 99%, MW 252.3), ranitidine hydrochloride ($C_{13}H_{22}N_4O_3S \cdot HCl$, >99%, MW 350.9), theophylline ($C_7H_8N_4O_2$, >99%, MW 180.2) and the internal standard, phenacetin ($C_{10}H_{13}NO_2$, MW 179.2), were purchased from Sigma-Aldrich, Inc. (St. Louis, MO). Their chemical structures are shown in **Figure 2-1**. HPLC grade water and methanol were acquired from Fisher Scientific (Fair Lawn, NJ). Pooled human plasma was obtained from LifeBlood Biological Services (Memphis, TN). All other materials were purchased from standard vendors and were of the highest available quality.

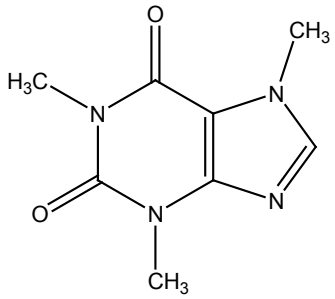
Instrumentation

The LC system consisted of a Shimadzu high performance liquid chromatographic system (Shimadzu Scientific Instruments, Norcross, GA, USA), coupled with a HTC PAL autosampler (Leap Technologies, CTC Analytics, Carrboro, NC). Chromatographic separation of acetaminophen, caffeine, phenytoin, ranitidine, theophylline, and the internal standard was performed on a Phenomenex Luna® 3 µm C18(2) column (50 mm x 2.00 mm, Phenomenex, Torrance, CA) with a gradient elution using mixtures of water and methanol, mobile phase A (95:5, v/v) and mobile phase B (10:90, v/v), both containing 0.05% formic acid. The optimum separation was achieved by increasing mobile phase B from 0% to 80% in the time period of 0 to 3 minutes, staying at 80% B from 3 to 5 minutes, and then dropping to 0 % B from 5 to 6 minutes. The flow rate was 0.3 mL/min. Detection was performed using a MDS Sciex API 3000 triple quadrupole mass spectrometer (Applied Biosystems, Foster City, CA) that was operated in positive ion mode with turbo electrospray ionization. All analyses were performed in the multiple reaction monitoring (MRM) mode. Instrument control and data acquisition were performed using the Analyst v1.4.2 software package (Applied Biosystems, Foster City, CA).

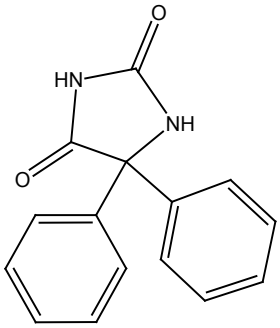
Optimization of the detection conditions was performed by direct infusion of the analytes (1 µg/mL, dissolved in methanol) from a syringe pump into the mass spectrometer. The auto tuning function of the Analyst software was used, and the



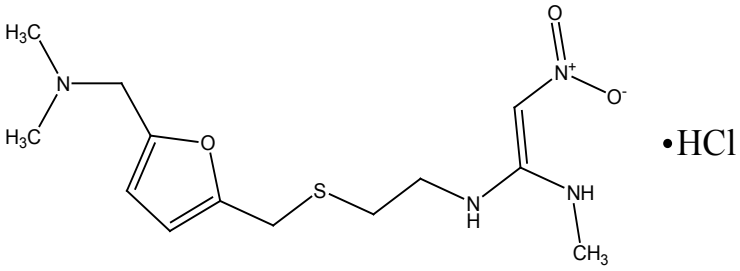
Acetaminophen $C_8H_9NO_2$ MW 151.2



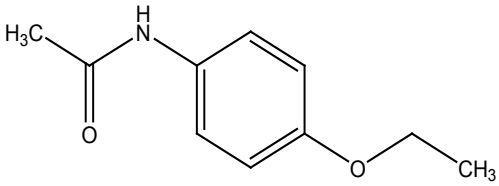
Caffeine $C_8H_{10}N_4O_2$ MW 194.2



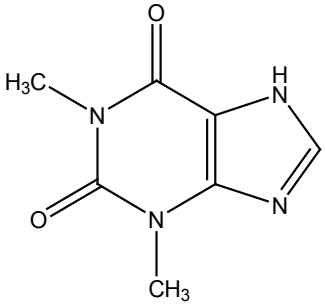
Phenytoin $C_{15}H_{12}N_2O_2$ MW 252.3



Ranitidine $C_{13}H_{22}N_4O_3S$• HCl MW 350.9



Phenacetin, internal standard $C_{10}H_{13}NO_2$ MW 179.2



Theophylline $C_7H_8N_4O_2$ MW 180.2

Figure 2-1. Chemical structures of analytes and internal standard

optimized parameters were used for the simultaneous detection of acetaminophen, caffeine, phenytoin, ranitidine, and theophylline. The parameter settings were as follows: turbo ionspray gas 7 L/min, nebulizer (nitrogen) gas 4.00 psi, curtain gas 8.00 psi, collision-activated dissociation gas 6.00 psi, ionspray voltage 5500 V, temperature 400 ºC, declustering potential 60 V, focusing potential 200 V, entrance potentials 10 V, collision energy 30 V, and collision cell exit potential 12 V.

Calibration standards and quality control samples

Primary stock solutions of analytes and IS were prepared at 1 mg/mL in methanol and stored at 20 ºC. 200 µg/mL working solution was prepared by combining equal volumes of acetaminophen, theophylline, caffeine, phenytoin, and ranitidine stocks. The highest calibrator at a concentration of 25,000 ng/mL was prepared by adding 125 µL of 200 µg/mL working solution into 875 µL of blank human plasma. Serial 1:2 dilutions of the highest calibrator in blank human plasma was used to produce 12 standard calibration samples with concentrations of 12.2, 24.4, 48.8, 97.7, 195.3, 390.6, 781.3, 1,562.5, 3,125, 6,250, 12,500, and 25,000 ng/mL. Internal standard working solution was diluted to 10 µg/mL in methanol. 1 mg/mL quality control (QC) standard solutions were prepared separately. Quality controls were prepared by adding small volumes of stock solutions to blank plasma. Three quality control levels at 100, 1,000, and 10,000 ng/mL were prepared and utilized for all drugs. Calibrators and controls were freshly prepared before each analysis.

Sample preparation

Sample preparation was performed by protein precipitation with methanol. 50 µL aliquots of plasma from calibration samples, quality control samples, or plasma specimens with unknown drug concentrations were transferred to 0.5 mL microcentrifuge tubes. 175 µL of ice-cold methanol containing 10 µL of the internal standard (10 µg/mL) was added to each tube. Samples were vortex-mixed briefly at high speed and kept on ice for 40 minutes. The samples were then centrifuged at 14,000 g for 10 minutes at 4 ^{o}C. Approximately 120 µL of the supernatant of each tube was transferred to an amber clean autosampler vial with insert for analysis. 10 µL of the aliquot solution was subsequently injected into the LC-MS/MS system.

Sample quantification

Concentrations of each analyte were determined based on the ratio of the peak area for their monitored mass transition and the peak area of the mass transition characteristic for the internal standard. A calibration curve covering the entire therapeutically used plasma concentration range was established for each analyte using linear regression analysis of the ratio of analyte peak area/internal standard peak area versus analyte concentration with a weighting factor of 1/x. Unknown analyte

concentrations were calculated from the calibration curve based on the measured peak area ratios for the various analytes monitored.

Validation

The developed LC-MS/MS assay was validated for linearity, accuracy, precision and recovery [105].

Linearity

Linearity was evaluated over the concentration range of 12.2 to 25,000 ng/mL for all analytes. Calibration standards were prepared freshly in duplicate for three validation runs on three separate days. The assay acceptance criterion for each standard concentration was ± 15% deviation of the nominal concentration, except for the lower limit of quantification, where a deviation of ± 20% was accepted.

Precision and accuracy

Precision was expressed as the percent relative standard deviation (%RSD) and accuracy was expressed as percent error [18]. The intra-day and inter-day accuracies and precisions of the assay were assessed by analyzing QC samples at three concentration levels (100, 1,000, 10,000 ng/mL). Five replicates of each QC sample were analyzed in the same batch and %RSD and percent error were calculated for each set of replicates per batch to determine the intra-day accuracy and precision. This process was performed three times over three consecutive days and %RSD and percent error were calculated for all 15 replicates per QC sample in order to determine the inter-day accuracy and precision.

Recovery and matrix effect

Recovery and matrix effect were assessed at three concentration levels (100, 1,000, and 10,000 ng/mL) for each of the analytes, comparing the peak areas of five replicates at each concentration for analyte standards in methanol and standards spiked before and after protein precipitation in human plasma [106-108]. Relative recovery was expressed as the ratio of the mean peak area of an analyte spiked before extraction to the mean peak area of the same analyte spiked post extraction in the same matrix multiplied by 100. Absolute recovery was calculated as the ratio of the mean peak area of an analyte spiked before extraction to the mean peak area of the same analyte spiked in methanol at the same concentration multiplied by 100. The matrix effect was evaluated by comparing the mean peak area of analyte spiked post extraction to the mean peak area of an equivalent concentration of the same analyte standard in methanol.

Results and Discussion

Method optimization

The assay development to simultaneously quantify acetaminophen, caffeine, phenytoin, ranitidine, and theophylline in small volume plasma specimens included optimization of the MS/MS detection, the chromatographic separation and the sample preparation procedures.

Based on the chemical structures of the analytes, an electrospray ionization interface (ESI) was used for ion generation. A Q1 full scan of each analyte and IS was acquired in both positive and negative mode when tuned under constant infusion at 600 µL/h of a 1 µg/mL methanol solution of the analytes. The signal-to-noise ratio was used as the measure of sensitivity [109]. The positive ion mode of the ESI was selected for all analytes and IS due to a greater sensitivity compared to the negative ion mode. The protonated form of the analyte molecules $[M + H]^+$ was monitored at m/z 152.2, 195.2, 253.3, 315.2, 181.3, 180.3 for acetaminophen, caffeine, phenytoin, ranitidine, theophylline, and IS, respectively. Similarly, the most abundant product ion of each analyte or IS was selected for observation in the multiple reaction monitoring (MRM) scan. The mass transitions selected for quantitative analysis were m/z 152.2 to 110.2 for acetaminophen, m/z 195.2 to 138.3 for caffeine, m/z 253.3 to182.3 for phenytoin, m/z 315.2 to 176.2 for ranitidine, m/z 181.3 to 124.0 for theophylline, and m/z 180.3 to 138.3 for phenacetin as IS. **Figure 2-2** depicts the product ion scan spectra of each analyte and the IS.

Due to the wide range in polarity of the five analytes, a single isocratic elution on a C18 column did not result in chromatographic separation within an acceptable run time. After evaluation of a variety of elution conditions, the separation, sensitivity, peak shapes and retention time were found to be satisfactory when using a gradient elution with a mobile phase of water and methanol containing 0.05% formic acid. All analytes and the IS had retention times of less than 6 minutes and the total assay run time was 8 minutes including the solvent equilibration time. **Figure 2-3** shows a representative chromatogram for a methanol solution containing 500 ng/mL of each analyte.

Assay performance

A chromatogram acquired from a blank human plasma sample spiked with 500 ng/mL acetaminophen, caffeine, phenytoin, ranitidine, theophylline is shown in **Figure 2-4**. For all analytes, good linearity in the calibration curves was achieved with correlation coefficients of $R > 0.9985$, or coefficients of determination of $R^2 > 0.997$. **Figure 2-5** depicts calibration curves for each analyte.

For acetaminophen, phenytoin, and ranitidine, the assay allowed quantification in a range of 12.2 to 25,000 ng/mL, for theophylline in the range of 24.4 to 25,000 ng/mL,

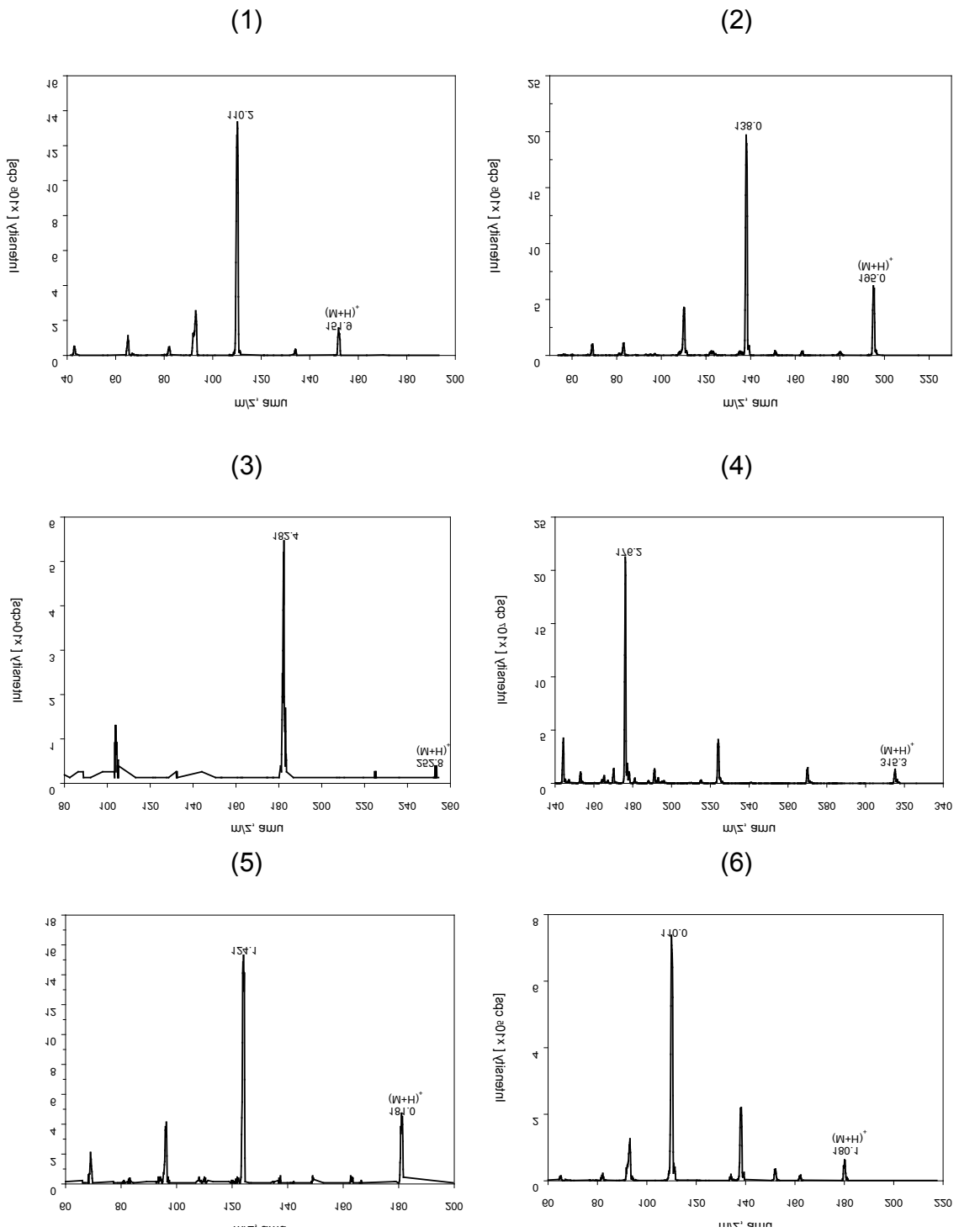


Figure 2-2. MS/MS product ion spectra of five analytes and internal standard

(1) acetaminophen, (2) caffeine, (3) phenytoin, (4) ranitidine, (5) theophylline, (6) phenacetin.

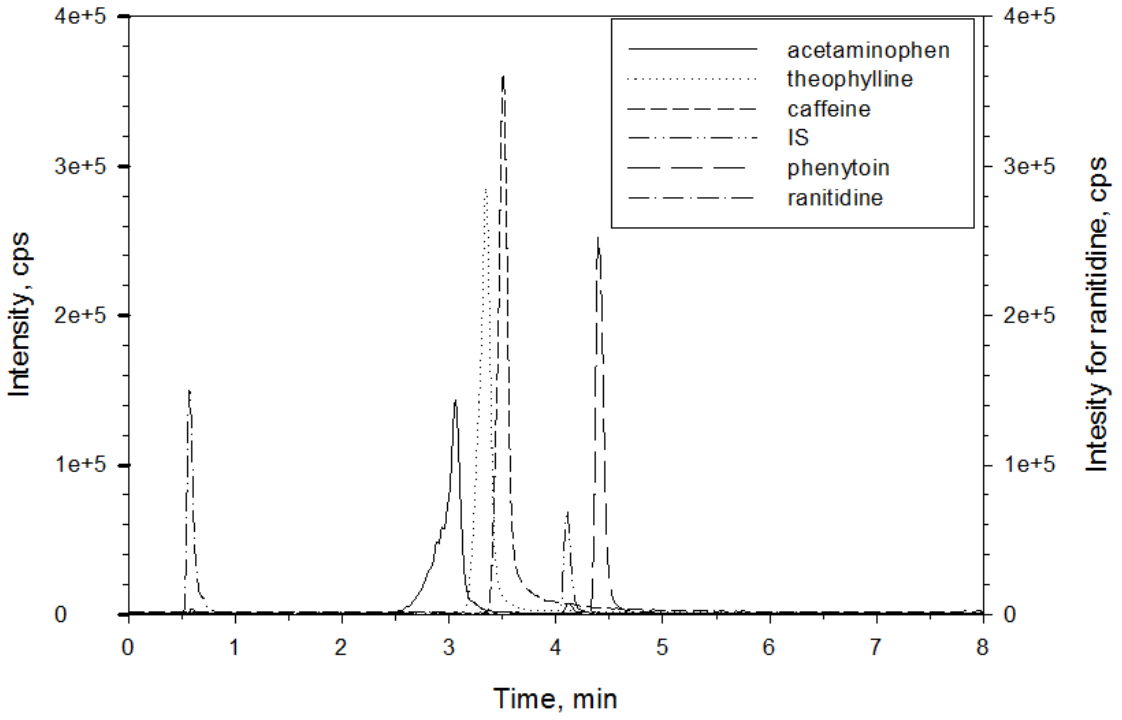


Figure 2-3. LC-MS/MS chromatograms acquired from a standard methanol solution containing 500 ng/mL analytes

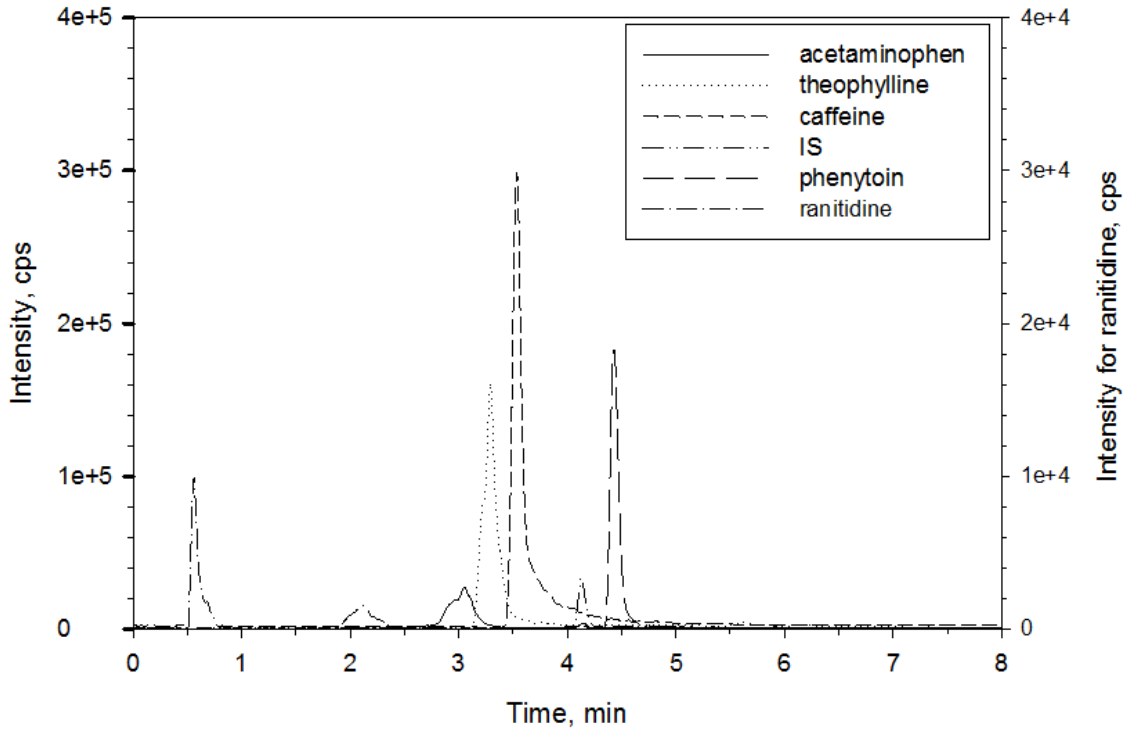


Figure 2-4. LC-MS/MS chromatograms acquired from blank human plasma spiked with 500 ng/mL of analyte drugs

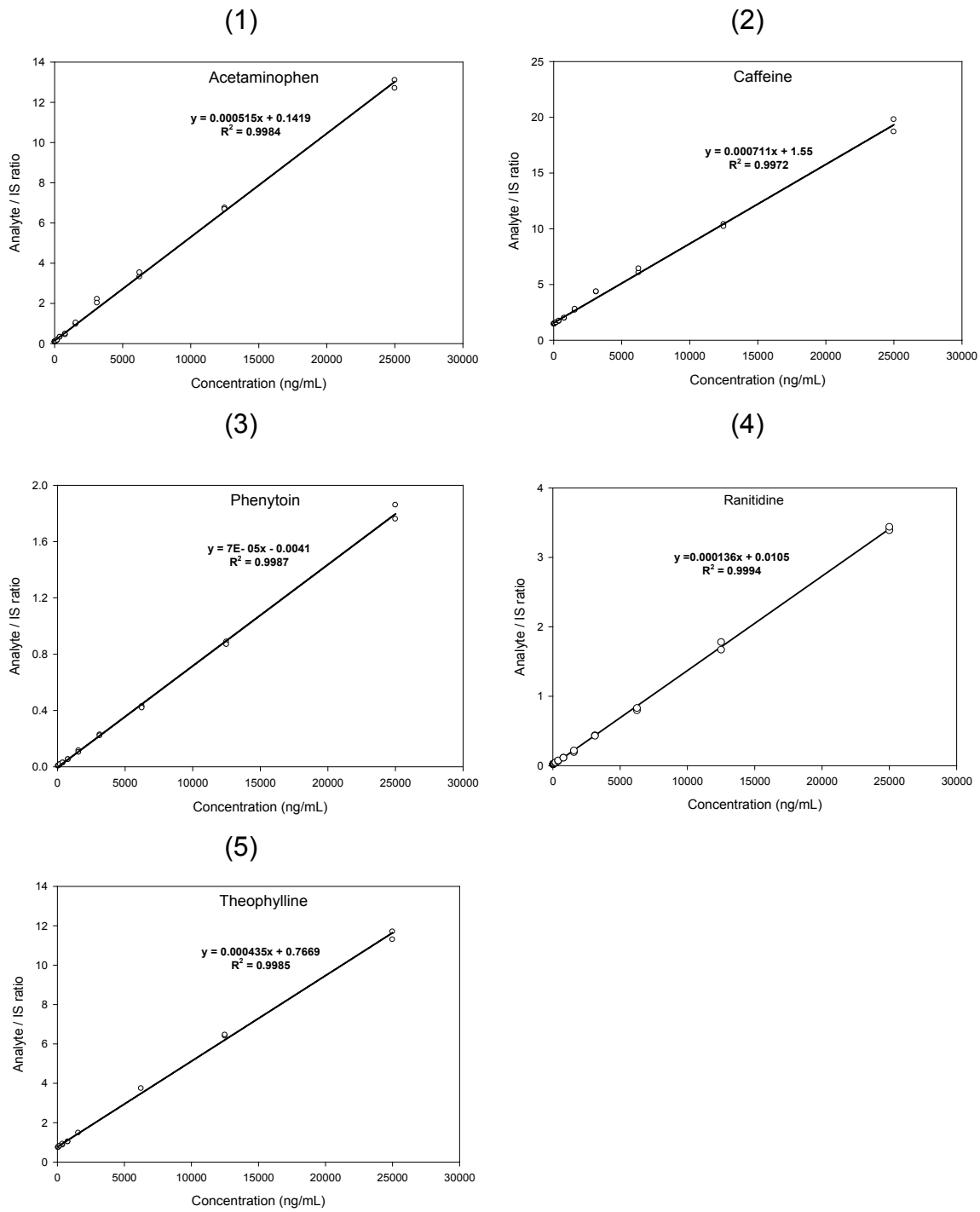


Figure 2-5. Calibration curves of analytes

(1) acetaminophen, (2) caffeine, (3) phenytoin, (4) ranitidine, (5) theophylline.

and for caffeine in the range of 48.8–25,000 ng/mL. These ranges cover each drug's therapeutically used concentrations in the neonatal age group described as 4,000–20,000 for acetaminophen [110, 111], 5,000–12,000 ng/mL for theophylline [112, 113], 8,000–20,000 ng/mL for caffeine [113], 6,000–15,000 ng/mL for phenytoin [103, 114], and 100–2,000 ng/mL for ranitidine [115-117].

The lower limit of quantification for each analyte was defined as the lowest concentration on the calibration curve with the signal-to-noise ratio (S/N) > 10 and is listed in **Table 2-1**. The upper limit of quantification was defined as the highest concentration on the calibration curve. Precision and accuracy for each analyte are summarized in **Table 2-2**. The mean accuracy ranged from 87.5 to 115.0% and the intra-day and inter-day precision was between 2.8–11.8% and 4.5–13.5%, respectively.

As some of the clinically measured concentrations may exceed the upper limits of quantification, a sample dilution procedure was also evaluated. The dilution procedure was conducted in five replicates for acetaminophen, theophylline, caffeine and phenytoin by using one half (25 µL) of the standard sample volume of plasma spiked with 50,000 ng/mL and 10,000 ng/mL of the analytes. All samples were diluted to 50 µL with blank plasma and underwent the same sample processing procedure as previously described. The intra-batch (within batch) accuracy and precision for the 1-to-2 dilution at both concentration levels ranged from 89.8–110.9% and 1.6–10.4%, respectively (**Table 2-3**), indicating that this dilution procedure can be applied to samples with very high analyte concentrations.

Table 2-4 presents the summarized data for absolute recovery, relative recovery and matrix effect. Since trace amounts of caffeine and theophylline were detected in all blank plasma batches available to us, the matrix effect and absolute recovery for the low and medium concentration range were not evaluated for these drugs. There was no significant interference detected from the plasma for any of the other analytes or the internal standard. Relative recoveries ranged from 85.6-118.3%, absolute recoveries ranged from 67.3–103.5%, and matrix effect assessments ranged from 61.7–112.0% for all of the analytes and the IS except ranitidine. Ranitidine showed a range of 105.4–118.5% for relative recovery, 26.3–41.9% for absolute recovery, and 23.5–35.3% for matrix effect. The high relative recovery suggests a good extraction efficiency of the protein precipitation method for all analytes and the IS. Similarly, percentage values for absolute recovery and matrix effect assessment were relatively high for all analytes except ranitidine, suggesting only a minor effect on the signal intensity by ion suppression from the matrix. Although ranitidine showed a relatively low absolute recovery and a pronounced matrix effect for human plasma, the analytical method was deemed acceptable for the intended purpose due to the satisfactory accuracy and precision obtained within the quantification range of ranitidine.

We also evaluated the effects of hemolysis, lipemia and hyperbilirubinemia (TBIL > 25 mg/dL and > 50 mg/dL) on the quantification of each drug at low (100 ng/mL), medium (1,000 ng/mL), and high concentration (10,000 ng/mL) levels. No interference in the analysis was noted when these factors existed separately or combined.

Table 2-1. Calibration range and lower limit of quantification (LLOQ) for each analyte

Analyte	Calibration Range (ng/mL)	LLOQ (ng/mL)
Acetaminophen	12.2 - 25000	12.2
Caffeine	48.8 - 25000	48.8
Phenytoin	12.2 - 25000	12.2
Ranitidine	12.2 - 25000	12.2
Theophylline	24.4 - 25000	24.4

Table 2-2. The accuracy and precision of the LC-MS/MS method for each analyte

Analyte	Nominal Concentration (ng/mL)	Accuracy (%)			Intra-day Precision (%RSD)	Inter-day Precision (%RSD)
		Day 1	Day 2	Day 3		
Acetaminophen	100	97.8	96.9	102.6	10.2	10.0
	1000	90.7	109.8	102.6	10.8	13.0
	10,000	87.5	108.8	94.6	5.3	10.9
Caffeine	100	104.3	107.9	104.2	6.9	6.8
	1000	95.8	103.5	100.7	11.8	11.5
	10,000	103.0	111.6	105.6	6.0	6.6
Phenytoin	100	90.6	105.8	105.6	10.7	13.2
	1000	93.9	115.0	99.8	4.1	9.2
	10,000	109.0	111.2	102.0	2.8	4.5
Ranitidine	100	112.4	113.3	99.6	4.5	7.1
	1000	106.4	110.4	105.3	5.7	5.6
	10,000	102.8	108.1	106	6.5	6.4
Theophylline	100	101.6	98.6	103.7	8.3	8.3
	1000	90.2	108.9	112.4	8.8	13.5
	10000	108.6	111.6	95.4	7.3	10.0

Table 2-3. Performance of a dilution procedure

Analyte	Dilution Factor	10000 ng/mL		50000 ng/mL	
		Accuracy	Precision	Accuracy	Precision
Acetaminophen	1 to 2	107.5	7.1	89.8	2.6
Caffeine	1 to 2	110.9	8.7	93.1	2.0
Phenytoin	1 to 2	109.6	4.7	107.3	1.6
Theophylline	1 to 2	97.9	10.4	103.7	2.0

Table 2-4. Recovery and matrix effect

Analyte	Relative Recovery (%)	Matrix Effect (%)	Absolute Recovery (%)
Acetaminophen	109.1-116.6	61.7-88.8	67.3-103.5
Caffeine	107.0-112.8	89.4*	96.2*
Phenytoin	92.5-109.6	72.7-110.5	68.5-102.2
Ranitidine	105.4-118.5	23.5-35.3	26.3-41.9
Theophylline	85.6-118.3	80.7*	87.1*
IS (Phenacetin)	94.3-97.9	100.0-112.0	94.2-102.0

* Since traces of caffeine and theophylline were detectable in blank plasma, matrix effect and absolute recovery were evaluated only at the highest concentration level for these analytes.

In addition, we investigated the potential interference of high concentration differences for analytes that are incompletely separated in the chromatographic process, particularly acetaminophen, caffeine and theophylline. Tested molar ratios of up to 13:1 did not result in any interference.

Figure 2-6 shows a representative LC-MS/MS chromatogram acquired from the plasma sample of a human subject taking four of the drugs captured by our assay. The analyzed concentrations were acetaminophen 7,010 ng/mL, theophylline 4,570 ng/mL, caffeine 401 ng/mL, and ranitidine 432 ng/mL.

Conclusions

In summary, we developed a rapid, accurate, sensitive, and reliable LC-MS/MS method to simultaneously quantify five drugs frequently used in the pharmacotherapy of preterm neonates. The analyte quantification can be performed from small volume human plasma specimens of only 50 μL, thereby facilitating an efficient use of limited blood samples in pediatric patients. This bioanalytical assay is highly useful in supporting clinical pharmacokinetic studies of these drugs in premature infants when combined with population-based modeling and simulation techniques [9].

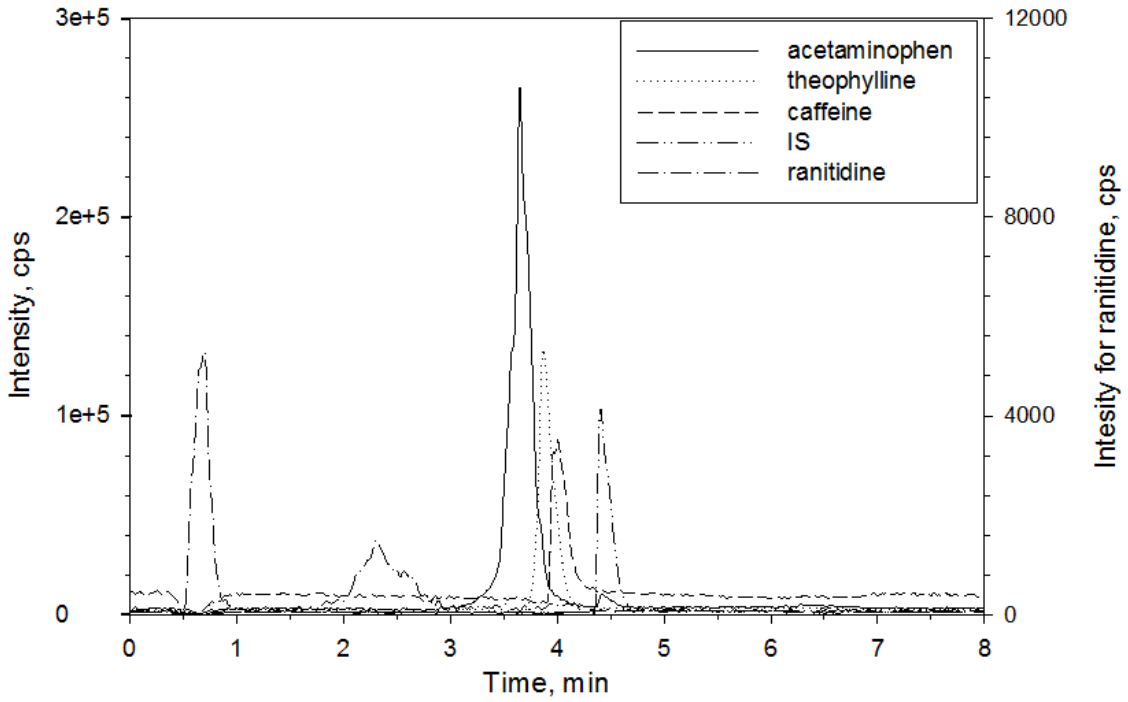


Figure 2-6. A representative LC-MS/MS chromatogram acquired from a subject's plasma

(1) acetaminophen 7,010 ng/mL, (2) theophylline 4,570 ng/mL, (3) caffeine 401 ng/mL, (4) ranitidine 432 ng/mL.

CHAPTER 3. SIMULATION-BASED SAMPLE SIZE OPTIMIZATION TO SUPPORT THEOPHYLLINE POPULATION PHARMACOKINETIC STUDY DESIGN IN PREMATURE NEONATES

Introduction

In Chapter 1, we discussed the major challenges in conducting pharmacokinetic (PK) studies in a pediatric population. Generally, ethical and practical concerns hamper clinical studies in premature infants as compared to adults and older children. The core problem is a lack of sufficient data for PK analysis due to the limited number of blood samples available per patient and the limited number of available patients. Chapter 2 describes an accurate and sensitive LC-MS/MS assay that was developed and validated for PK studies in premature infants. The presented bioanalytical assay allows for simultaneous quantification of five frequently used drugs from a single plasma sample as small as 50 μL. Therefore, this assay provides an opportunity for clinicians to assess the PK of multiple drugs in premature infants simultaneously, which is valuable in expanding our knowledge of PK in this population. Modeling and simulation is another effective tool in pediatric PK studies. It allows researchers to explore "what if" scenarios and therefore facilitates the optimization of study design in drug development and applied pharmacotherapy in terms of sampling scheme (the time and number of blood samples collected per patient) and sample size estimation (minimal number of subjects needed under the given sampling scheme), resulting in a reduced need for experiments and invasive study procedures. In this chapter, a full model-based, optimal-sample-size estimation for a population pharmacokinetic (PopPK) study is presented. Theophylline was used as a model drug in the current study. The results may be applied in future PopPK studies in premature infants with drugs commonly used in the neonatal intensive care unit (NICU).

Specific sampling design in pediatric studies

The aim of a pediatric study to characterize the PK disposition of drugs commonly used in preterm infants is to ultimately use this information to guide an individualized dosing strategy. To minimize the number of needed patients and the procedures needed with each individual in this vulnerable population, an optimum clinical trial design is desirable. However, a well-designed clinical study with an optimal preplanned blood sampling schedule may not be practical and ethical in the NICU setting, since research in this study population can only be performed within the context of therapeutically necessary interventions, thereby limiting the available volume, frequency and timing of PK blood sampling. Thus, the data collected from routine therapeutic drug monitoring (TDM) have been suggested as an alternative for a PK study [118, 119]. One major restriction using the data from routine TDM is that most concentration measurements are from trough levels, therefore lacking information for the estimation of volume of distribution [120], which mainly determines the loading dose of a drug treatment. Since a small quantity of (50 μL) plasma is sufficient for an assay quantitation

of drug concentrations, the open question is whether an aliquot from leftover routinely obtained blood samples can be utilized in PK studies. A NICU patient usually undergoes frequent routine laboratory assessments for complete blood count and biochemical tests. In some critical situations, blood gas, electrolytes, and blood glucose are also monitored as frequently as every one hour for the purpose of therapeutic guidance. We thus hypothesized that any available leftover samples from those diagnostic blood samples plus TDM data could be used for PK evaluations. One of the advantages of this study design, which uses an opportunistic sampling approach as opposed to a more traditional preplanned PK approach, is that no extra blood draw will need to be imposed on the patients, therefore minimizing patient risk. Meanwhile the potential gain in knowledge of specific drug disposition will significantly extend our understanding of drug therapy in extremely premature infants. It is expected that over the course of the study, the sampling times are variable among the patients and the number of concentration measurements per subject will vary as well due to the different clinical requirements for blood specimens. Therefore, both sparse and dense samplings with random sampling time allocations for different drugs are expected in this study. In this particular case, we asked the question how many patients are minimally required for a PopPK study to ensure accurate estimation of the relevant model parameters to reliably detect clinically meaningful differences.

Sample size and population pharmacokinetics

Population pharmacokinetics (nonlinear mixed effect modeling)

The population pharmacokinetic (PopPK) approach, using nonlinear mixed effects modeling, allows for the simultaneous analysis of pooled data from multiple patients and provides population-typical as well as individual PK parameter estimates. One popular tool for PopPK analyses is the NONMEM® software. The term PopPK is used synonymously with nonlinear mixed effects modeling today. It has been frequently applied in pediatric studies due to the ability to extract information from sparse and unbalanced sampling data [11, 76, 121].

Sample size for PopPK

Inefficient sampling design and unsuitable sample size may lead to a failed population pharmacokinetic study [122]. Although it is always favorable to acquire parameter estimations from a large sample size, the key is to determine the minimum but adequate number of subjects needed to balance the study cost and duration, and to ensure a study with adequate power. Sample size has been demonstrated to be one critical determinant. Its calculation for analyses in nonlinear mixed effects modeling has not been clearly defined. A number of publications suggested sample size determinations for PopPK studies. Those proposed methodologies were either formulae based by using the Wald statistic with first-order linearization of the nonlinear mixed effects model [123, 124] or simulation based on the likelihood ratio test or a confidence interval (CI) approach [125, 126]. All of these studies determined the minimum sample size needed

for detecting some level of difference for a parameter between two subpopulation groups based on a prespecified hypothesis test, i.e., determining sample size as a function of statistical power and clinically meaningful effect size (together with a given type I error probability). For example, the hypothesis to be tested could be: is clearance in group one 15% lower than that in group two? How many subjects do I need to detect this difference? How many subjects do I need if the difference in clearance is 30% or 40%? Usually only the primary model might be involved, and some categorical covariates (such as gender groups, two or more age groups) might be investigated as a subgroup.

In contrast, our study focused on two issues. The first was to estimate the typical population pharmacokinetic parameters and their between-subject variability with certain levels of accuracy and precision. Secondly, we wanted to assess and identify potential covariates, that are significant predictors of pharmacokinetic parameters, such as CL and volume of distribution [72]. Accordingly, the major concerns of sample size determination in such studies should not focus on testing hypothesis or detecting the differences in parameters between subgroups. Instead, sample size estimations would be carried out for the purpose of a successful PopPK study: how many subjects should be recruited so as to (1) obtain parameter estimations in the model with adequate accuracy and precision; (2) reliably determine covariate effects by separating a covariate model from the base model or its nested covariate model. Ogungbenro and Aarons [127] proposed a confidence interval approach for the sample size determination of a PopPK study when there is no clear hypothesis to be tested. In the current study, we attempted to extend the application of this approach to explore the impact of sample size on the quantification of continuous covariate effects (weight and age) in premature infants. In our study, given a more complicated study design in premature infants, time-dependent covariates and randomized sampling design were included; dose levels, sampling times and number of concentrations per subjects were different for different patients. We here present a full model-based sample size estimation using a simulation approach.

Objective

This simulation study was to explore sample size requirements for observational PopPK studies in premature infants using theophylline as a model drug. It took prior information (model, parameter estimates, between-subject variability and residual variability) to identify a sample size that could provide unbiased and precise estimates for the fixed and random effect parameters and covariate effect determination.

Methods

Overview of methodology

Figure 3-1 illustrates how the simulations were conducted to assess different sample sizes for their ability to yield meaningful results given the same PopPK study

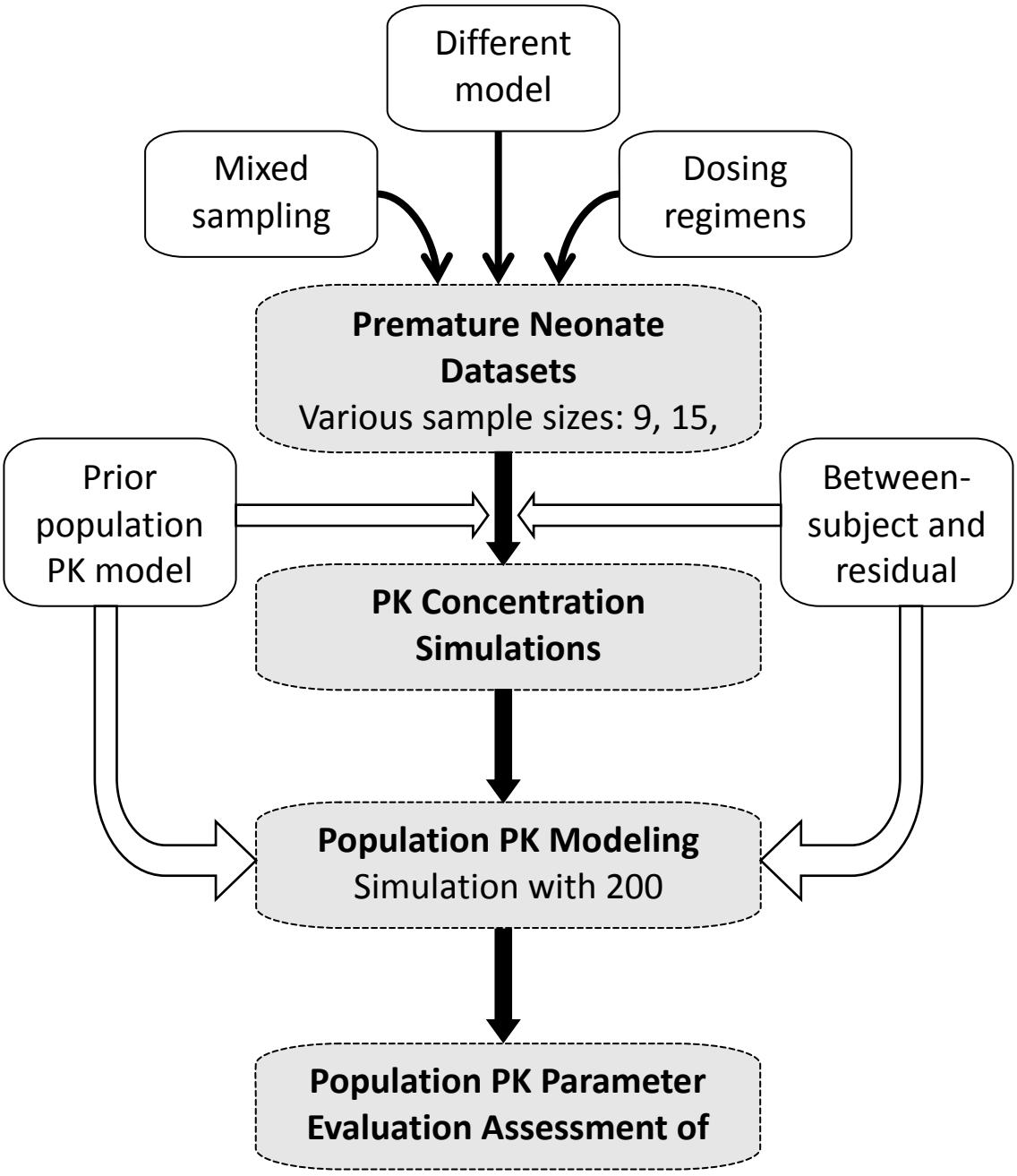


Figure 3-1. Overview of the methodology of the simulation and estimation steps

This figure details how the simulations were conducted to assess the designs and their results.

design. Main steps included:

- SAS 9.1 was used to generate the datasets, including covariates and dosing history.
- 200 independent datasets for each sample size (9-200 subjects per study) were simulated.
- A previously established population pharmacokinetic model and model parameters were employed to simulate all concentration datasets. The step was accomplished by using NONMEM version VI, Level 2.0 (ICON Development Solutions, Ellicott City, Maryland) with GNU Fortran 77 (g77) version 2.95 (Free Software Foundation, Cambridge, Massachusetts).
- The PopPK model used for concentration simulation was applied to perform re-estimation. Parameter estimates were then obtained for each dataset.
- The process was repeated with 200 independent datasets. Median and 95% confidence intervals (CIs) were then obtained for each parameter from 200 estimates.
- The bias (accuracy) and precision in the estimates of the population mean PK parameters and variance components were evaluated.
- The impact of sample size on the parameter estimations in the model was evaluated based on the proportion of the total number of simulations where the estimated parameter values would fall inside a pre-specified interval considered "close enough" to their respective true value.
- Power to detect two continuous covariates, body weight and postnatal age, was calculated as a fraction of tested statistics obtained in total number of simulations.

From the literature review, drugs of interest, such as caffeine, theophylline, acetaminophen and phenytoin, all exhibited one compartment disposition model, with age and weight being the two most important covariates of clearance and volume of distribution [85, 128-130]. Therefore, a one-compartment model was considered to examine the influence of sample sizes on the parameter estimation; weight and age were used to illustrate the covariate effect determination in the PopPK study.

Datasets and simulating study design

Simulated datasets containing demographic data, dosing histories and sampling times were generated using SAS version 9.1 (SAS Institute, Cary, North Carolina).

Covariate simulation: demographic variables

Demographic variables, including birth weight (BW) and postnatal age (PNA) at the entry time (PNA0), were generated by randomized uniform distribution based on defined ranges in published data [131]. Postnatal age ranged between 1 and 25 days at entry time into the study. Birth weight ranged from 400 g to 1500 g. Considering the fact that premature neonates might be born at different gestational ages, no correlation was given between PNA0 and BW. PNA was then derived from PNA0, and body weight (WT) derived from BW and PNA according to biologically rationale development curves through the 14-day sampling window, using published growth behavior [132]. Thus

$$\text{PNA0 (days)} \sim U(1, 25) \qquad \text{BW (g)} \sim U(400, 1500)$$

Where U (a, b) refers to a uniform distribution with lower (a) and upper (b) limits.

Dosing history

All subjects were simulated to receive orally administered theophylline with a loading dose of 6 mg/kg followed by maintenance doses of 3 mg/kg every 12 hours for 14 days.

Pharmacokinetic samplings

The simulations assumed that future population studies in premature infants will be conducted with an opportunistic sampling approach as opposed to a more traditional pharmacokinetic approach. A mixed and unbalanced sampling design (both sparse and rich sampling) was included. All sampling times were simulated as random occurrences rather than a pre-planned sampling scheme as only blood samples drawn for clinical purposes would be used in the study. For each design (sample size), one third of individuals supplied 2 sampling measurements, one third supplied 4 measurements, and the remaining one third provided 8 sampling points.

Simulation scenarios

Sample sizes of 9, 15, 20, 40, 60, 80, 100, and 200 patients per study were simulated. For each sample size, 200 independent datasets were generated under the same condition.

Pharmacokinetic model and statistical model

Theophylline was used as a model drug to illustrate the method described in this section. A previously established population pharmacokinetic model with model parameters for both fixed and random effects was used in the evaluation [131]. The PK model was described by a one compartment model with first order absorption and first order elimination. The population pharmacokinetic model, parameterized in terms of

clearance (CL), volume of distribution (V) and bioavailability (F1), was as follows:

$$TVCL\ (mL/hr) = \theta_1 * WT\ (g) + \theta_2 * PNA\ (days)$$

$$TVV\ (L) = \theta_3 * WT\ (g)$$

$$TVF1 = \theta_4$$

Where TVCL is the population typical value of clearance, TVV is the population typical value for volume of distribution, and TVF1 is the typical population value for bioavailability. θ_1 is the coefficient for the effect of WT on CL. θ_2 is the coefficient for the effect of PNA on CL and θ_3 is the coefficient for the effect of WT on V. Following literature, we assume these parameters to be θ_1= 0.0123, θ_2= 0.377, θ_3= 0.000937 and θ_4 = 0.918 [131].

The between-subject variability (BSV) of the population typical value of CL and V was expressed by proportional error models as follows:

$$CL_j = TVCL * (1+\eta_{CL,j})$$

$$V_j = TVV * (1+\eta_{V,j})$$

$$\eta_{CL,j}, \eta_{V,j} \sim N\ (0, \omega^2)$$

Where CL_j and V_j are parameter estimations for the j^{th} individual in the study. $\eta_{CL,j}$ and $\eta_{V,j}$ represent random variables normally distributed with zero means and variances of ω_{CL}^2 and ω_V^2, respectively. ω_{CL}^2 and ω_V^2 can be estimated by NONMEM and represent the between-subject variability in the population.

Residual variability was described by an additive error model as follows:

$$C_{obs,ij} = C_{pred,ij} + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N\ (0, \sigma^2)$$

Where $C_{obs,ij}$ is the i^{th} observed concentration in the j^{th} subject, $C_{pred,ij}$ is the i^{th} model predict concentration in the j^{th} subject and ε_{ij} is the deviation of $C_{obs,ij}$ from $C_{pred,ij}$. ε_{ij} is a normally distributed random variable with an average value of 0 and variance of σ^2.

Simulation values of between-subject variability of CL and V in terms of coefficient of variation (CV %) were set at 15% on clearance and 43.5% on volume of distribution --the same as the estimates reported in the original publication [20]. The BSV on bioavailability F1 was set at 0. Residual variability in terms of standard deviation was set at 1.93 mg/L, representing the differences between the observed and predicted concentrations in the study population [131].

Simulating concentration profile

Using the pharmacokinetic model described above and all the parameters for both fixed and random effects, concentration profiles for theophylline in premature infants were simulated with nonlinear mixed effect modeling, as implemented in NONMEM VI, using the first order conditional estimation (FOCE) method.

Data Analysis

Parameter estimation

For each specified sample size, 200 replicates (or independent datasets) were analyzed. The model described earlier was fitted to all the datasets. For each replicate dataset, fixed effect parameters θ_1, θ_2, θ_3 and variability parameters ω_{CL}, ω_V and σ were re-estimated by NONMEM VI using the same model. Bioavailability parameter θ_4 was fixed since exclusive oral data was simulated in this study. First-order conditional estimation method was used in all cases throughout the study. Approximately 2-15% of estimation runs in each study experienced terminated minimization with reported parameter estimates; these model-fitting processes were repeated with adjusted initial estimates. This procedure did not necessarily lead to successful convergence. However, at least 90% of the 200 simulation runs converged successfully for each study. All reported parameter estimates, including those from terminated minimizations, were used in the analysis. This procedure was carried out in each study design, with a sample size of 9, 20, 40, 60, 80, 100, and 200 subjects per study.

The parameters obtained from the 200 simulation datasets for each sample size were compared with the numbers used in the concentration simulation step to assess the bias and precision in the estimates of the population mean PK parameters and variance components. Median parameter estimates were compared to the 'true' parameters of the originating model, and 95% CIs for 200 replicates of each study were determined and evaluated by visual inspection [122] to detect trends in the results. Mean prediction error (%MPE) and root mean square error (RMSE) were computed as indices of accuracy and precision using the following formula:

$$\%MPE = \frac{1}{n}\left[\sum_{j=1}^{n}\left(\frac{\hat{P}_{jk} - P_k}{P_k} * 100\right)\right]$$

$$RMSE = \left[\frac{1}{n}\sum_{j=1}^{n}\left(\hat{P}_{jk} - P_k\right)^2\right]^{\frac{1}{2}}$$

Where n = number of simulations (n = 200), $\hat{P}_{jk}$ is the value of parameter P_k estimated in the j[th] simulation and P_k is the criterion value of parameters, representing both fixed- and random-effect parameters.

In order to assess whether a certain number of subjects (sample size) is enough to capture the true value of the parameter P_k in the PopPK study under the given opportunistic sampling design (blood samples collected per subject), the propensity of the resulting parameter estimates to fall within pre-specified narrow intervals containing the true parameter values was also investigated for each sample size. Typically, the statistical power of a hypothesis test is defined as the probability that we correctly reject the null hypothesis when a certain minimal effect size (deviation of the true value from the value assumed under the null hypothesis) is indeed present (equivalent to 1- β = 1- probability of a Type II error). In a simulation study, statistical power is typically not determined by formula-based computations but determined by the percentage of correctly rejected null hypotheses in repeated computer experiments that emulate certain realistic study settings [133]. In the current study, we compare the estimated parameter $\hat{P}_{jk}$ to the criterion parameter from the literature, P_k, (the ideal ratio of $\hat{P}_{jk}$ to P_k is 1 if $\hat{P}_{jk}$ coincides with the true parameter value). If a difference > 20% (20% precision level) is determined to be significantly different, a range of 0.80-1.25 for the ratio is an acceptable criteria based on the two one-sided tests procedure [134]. That means if the ratio of $\hat{P}_{jk} / P_k$ falls in the range of [0.80, 1.25], we considered the true parameter could be accurately estimated at the 20% precision level (i.e., would be reproducible for practical purposes). Note that this procedure is somewhat similar to the traditional "power" concept but has important distinctions; most notably we compare the obtained ratio to a pre-specified interval of acceptable values, and hence do not compute a confidence interval that would vary in each iteration of the simulation as is the case in statistical hypothesis testing. Furthermore, we are interested in how often we estimate the parameter "close enough" to the true value when simulating from this true model and do not evaluate how often we reject a certain value when the true model indeed differs in a particular way as would be the case when determining traditional power of a study. We proceeded as follows: For each investigated sample size, 200 ratios were obtained from the 200 simulations. The relative success of our computer experiments ("power") were computed as the number of times we "correctly" identified the parameter value (estimate within the pre-specified values) divided by 200. Namely, the percentage of the 200 ratios that fell within the limits of 0.80-1.25 is referred to as the relative success ("power") of the study. This procedure was also repeated at precision levels of 30% and 40% for parameter estimations with the corresponding ranges of [0.70, 1.43] and [0.60, 1.67] for the $\hat{P}_{jk} / P_k$ ratios.

Covariates effect determination

The impact of sample size on covariate effect determination was investigated through comparison of objective function values (OFV) of a covariate model and its nested model in 200 replicate datasets. OFV is proportional to -2 log likelihood of the

data and is a global measure of goodness of fit. Postnatal age and body weight were selected as the predictors of clearance, and body weight was selected as the predictor of volume of distribution [131]. Model improvement when including one covariate was evaluated based on chi-square distribution with one degree of freedom. A statistically significant model improvement was associated with a decrease of OFV by 3.84 when P = 0.05. More stringent criteria commonly used are $\chi^2_{1, 0.01} = 6.64$ and $\chi^2_{1, 0.001} = 10.83$. To be conservative, estimation runs with failed convergence were repeated with adjusted initial estimates until successful minimization was gained in all 200 replicates. Model separation was based on the decrease of objective function values (ΔOFV) at 3 levels (i.e., P = 0.05, P = 0.01 and P = 0.001).

For each proposed sample size design, power was determined by calculating the fraction of simulations that ΔOFV achieved or exceeded the prespecified criteria for at the three significance levels.

Results

Dataset generation

Based on a reported longitudinal postnatal growth study in very low birth weight infants [132], average daily weight gain in g/day varied across 100-g birth weight intervals, ranging from 15.27 to 27.77 g/day. Gestational age, race and gender had no significant influence on the growth rate within each 100-g birth weight interval. So the average daily increments stratified by 100 g birth weight interval (**Table 3-1**) were employed to compute the body weight gain in the simulation.

An example for the mixed, unbalanced and randomized sampling design is shown in **Figure 3-2.** The representative study had included 9 simulated subjects with 42 concentrations; of them, 3 subjects contributed 2 concentration measurements, 3 subjects contributed 4 concentration measurements, and the other 3 subjects contributed 8 concentration measurements. The sampling allocation time ranged from 0.5 hr to 329 hr post loading dose within 14 days (336 hr) treatment. The sampling time spread between any two concentration measurements for each subject ranged from 2 hr to 237 hr, with a median of 32 hr. All sampling times were simulated as random occurrences over a 14-day sampling window, representing the paradigm of flexible blood sampling performed at any time. It should be noted that all datasets were generated independently so that the descriptive statistics for sampling time and demographic information, such as PNA and birth weight, varied in the simulations of the 200 different studies.

Precision and accuracy of parameter estimation

The mean parameter estimates of the 200 simulations using sample sizes of 9, 15, 20, 40, 60, 80, 100 and 200 subjects are presented in **Table 3-2**. Under the given study

Table 3-1. Average daily increments (g/day) used for body weight calculation stratified by birth weight interval

Birth Weight Interval (g)	Weight Gain (g/day)
≤ 600	15.27
601-700	16.81
701-800	18.6
801-900	20.06
901-1000	21.04
1001-1100	22.83
1101-1200	24.73
1201-1300	26.34
1301-1400	27.15
1401-1500	27.77

Source: Modified with permission. Ehrenkranz, R.A., et al., *Longitudinal growth of hospitalized very low birth weight infants.* Pediatrics, 1999. **104** (2 Pt 1): p. 280-9.

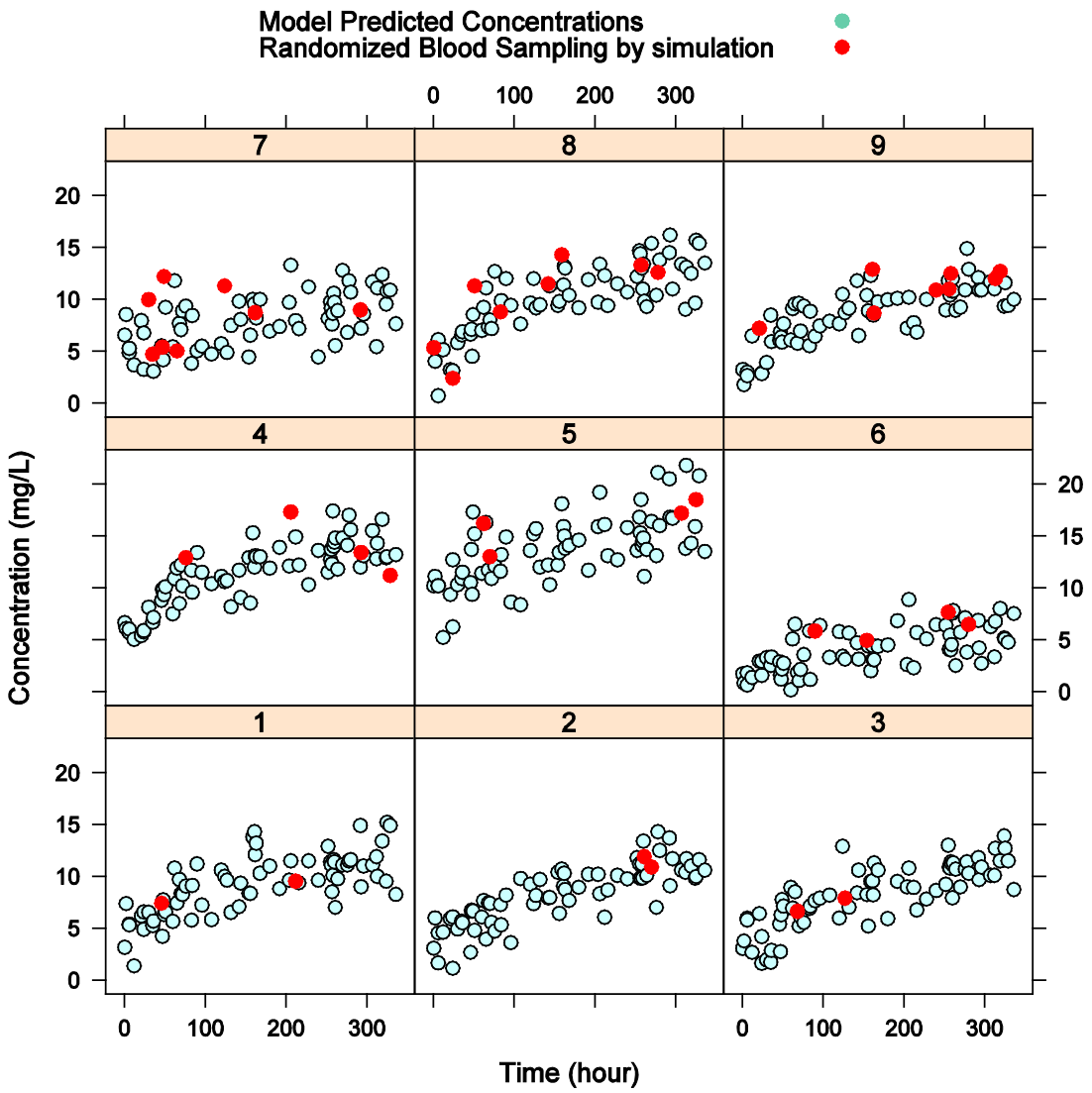


Figure 3-2. A representative study showing the mixed, unbalanced and randomized samplings from a 9-subject study

Table 3-2. Central tendency (median parameter) of estimates for the simulations at different sample size

Sample Size	θ1	θ2	θ3	ω_{CL}	ω_V	σ
Criterion Value	**0.0123**	**0.000377**	**0.000937**	**0.0226**	**0.189**	**3.72**
9	0.0120	0.000366	0.000909	0.0110	0.099	3.69
15	0.0124	0.000365	0.000931	0.0174	0.149	3.57
20	0.0121	0.000381	0.000911	0.0181	0.150	3.65
40	0.0126	0.000358	0.000931	0.0196	0.156	3.74
60	0.0128	0.000357	0.000909	0.0216	0.156	3.72
80	0.0126	0.000365	0.000927	0.0209	0.161	3.75
100	0.0126	0.000370	0.000912	0.0223	0.157	3.76
200	0.0127	0.000357	0.000918	0.0220	0.161	3.77

design, median parameters were well estimated across all sample size groups. When visually inspected, 95% CIs of PopPK parameters as measure of precision converged as sample size increased and became much narrower and remained stable when the sample size was > 60 (**Figure 3-3**).

The %MPE results are shown in **Table 3-3** and illustrated in **Figure 3-4.** A smaller %MPE value indicates an on average smaller (relative) deviation of the estimated parameter value from the true value (preferable). All sample sizes gave unbiased estimates for all parameters except for two variability parameters, BSV on clearance estimation (ω_{CL}) and BSV on volume of distribution estimation (ω_V). %MPE was below 5% for $\theta_{1\sim3,}$ and σ estimations at all investigated sample sizes except for one estimate of θ_2 showing as -6.2%. Substantial bias in variance of between-subject variability for V and CL (ω_{CL} ω_V) is noted for small sample sizes. %MPE as large as -33% for $\omega_{CL,}$ and -32% for ω_V was observed at sample size of 9; but dropped to -6.5% and -6.2%, respectively, when sample size increased to 40, indicating a marked increase in accuracy with increasing sample size. The bias for ω_{CL} and ω_V estimations was entirely negative, while both positive and negative biases were observed for other parameters in the model. Optimal sample size was evaluated by assuming a percentage coefficient of variation (CV%) at 15% for CL and 43.5% for V. Histograms of CV % for CL and V by different sample size are shown in **Figure 3-5** and **Figure 3-6.** The results suggest that a substantial improvement in the estimation of variance parameters is correlated with an increased sample size.

The results of RMSE for the simulations are presented in **Table 3-4,** and its CV% is illustrated in **Figure 3-7**. RMSE generally decreases in all parameter estimations as the sample size increases, indicating increasing precision with increasing sample size. For a given sample size, the estimations of fixed effect parameters are better than those of the random effect parameters. The number of subjects does not seem to have as significant an influence on the precision of parameter estimations as long as it reaches 40 for fixed effect parameters and 60 for covariance parameters.

The influence of the number of subjects on the relative success of the study as defined here ("power") was also investigated. Plots of relative success against sample size at various precision levels are presented in **Figure 3-8**. With our randomized sampling design, the relative success of our parameter estimation was deeply influenced by sample size, parameter of interest and the selected precision level. For example, assuming 20% difference was allowed in parameter estimation, a study would require 20 subjects to give in 80% of the cases "close enough" estimates for θ_1, while at least 100 subjects would be required to achieve the same performance for θ_2 under the current sampling and study design. A relatively high success rate (≥ 0.8) was shown in all sample sizes for 30% (ratio limit 0.70-1.43) and 40% (ratio limit 0.60-1.67) precision levels for θ_1, θ_3 and σ. The number of subjects required for θ_2 also dropped to 40 and 20, respectively, at those two levels. The success rate of estimating the BSV parameters ω_{CL} and ω_V was much lower compared to the other parameters (**Figure 3-9**). To obtain successful estimation with a probability > 0.6, 20, 40 and 200 subjects in each study were considered to be sufficient at a precision level of 40%, 30% and 20%, respectively.

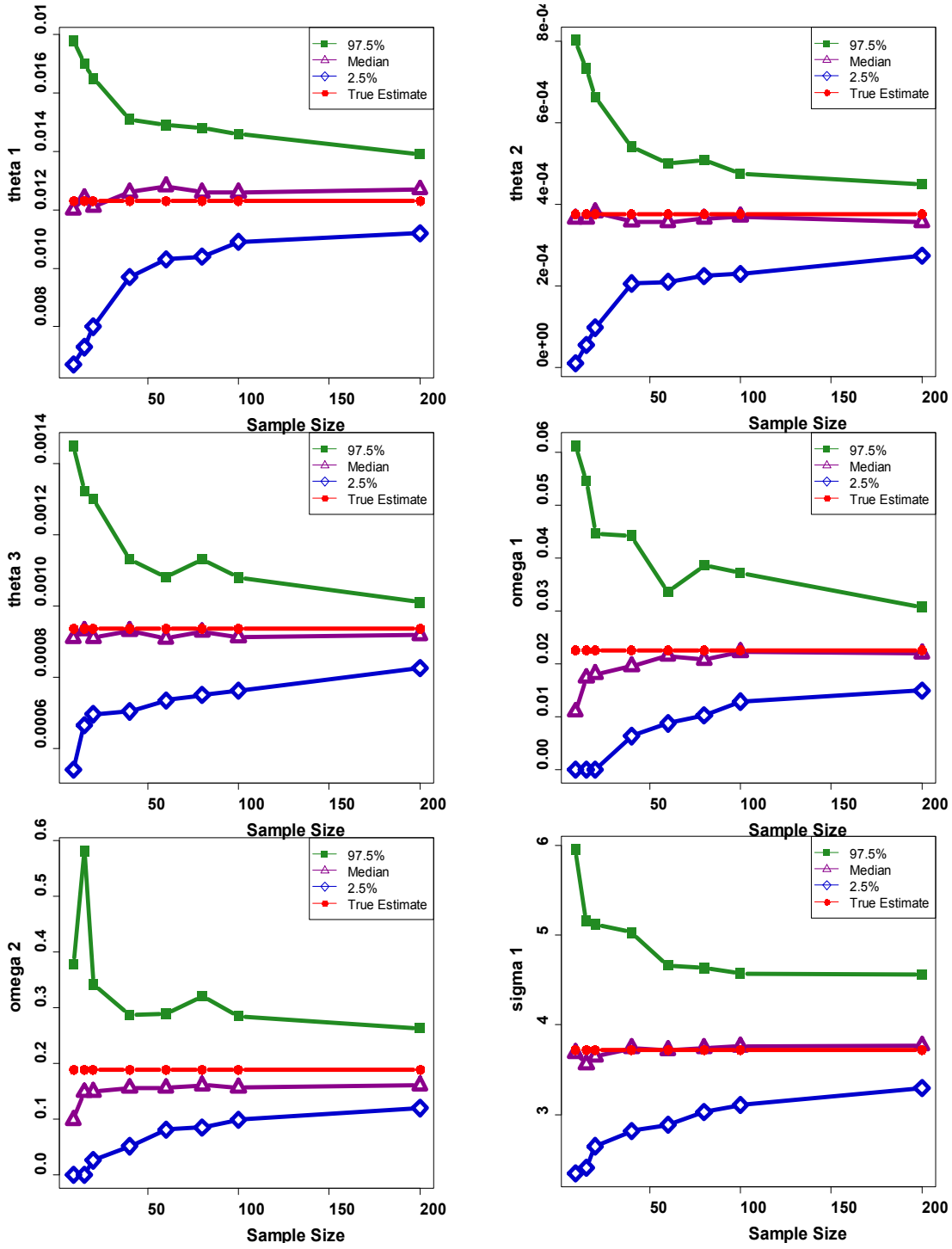


Figure 3-3. Median and 95% CIs for PopPK parameters and variance parameters from 200 simulated datasets

Table 3-3. %MPE of parameter estimates with various sample sizes

Sample Size	θ1	θ2	θ3	ω_{CL}	ω_V	σ
9	-1.884	4.265	0.684	-33.054	-32.119	1.036
15	1.153	-1.638	0.432	-19.287	-10.914	-1.813
20	-0.373	4.229	-0.584	-15.043	-13.471	-0.370
40	1.799	-2.621	-0.755	-6.486	-12.401	3.739
60	3.463	-6.229	-3.126	-2.591	-12.428	0.780
80	2.898	-3.789	-1.170	-2.525	-6.966	0.667
100	2.600	-2.385	-3.152	2.212	-9.565	1.816
200	3.172	-4.711	-1.767	-1.473	-7.221	2.829

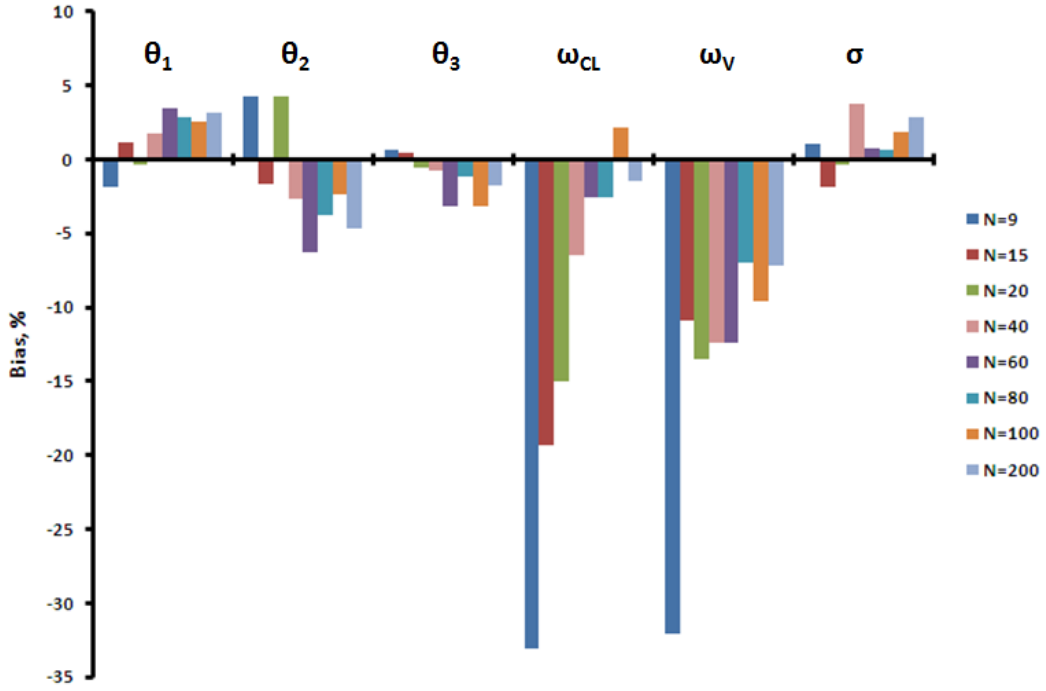


Figure 3-4. Accuracy of parameter estimates in different sample size groups

Bias is expressed in terms of %MPE.

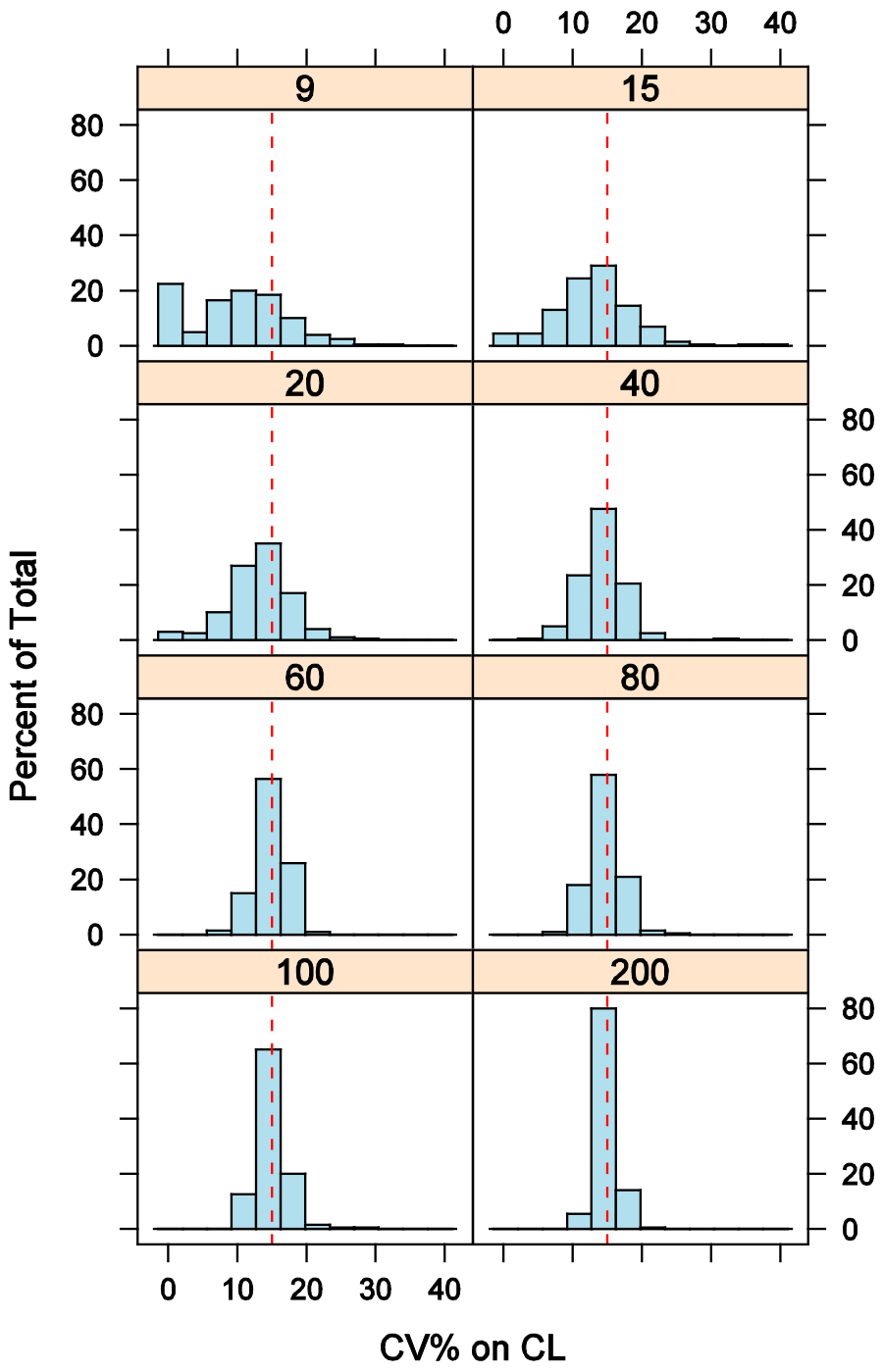


Figure 3-5. Estimated BSV on population CL in terms of CV% vs. numbers of subjects

The red dash line indicates the CV% = 15%.

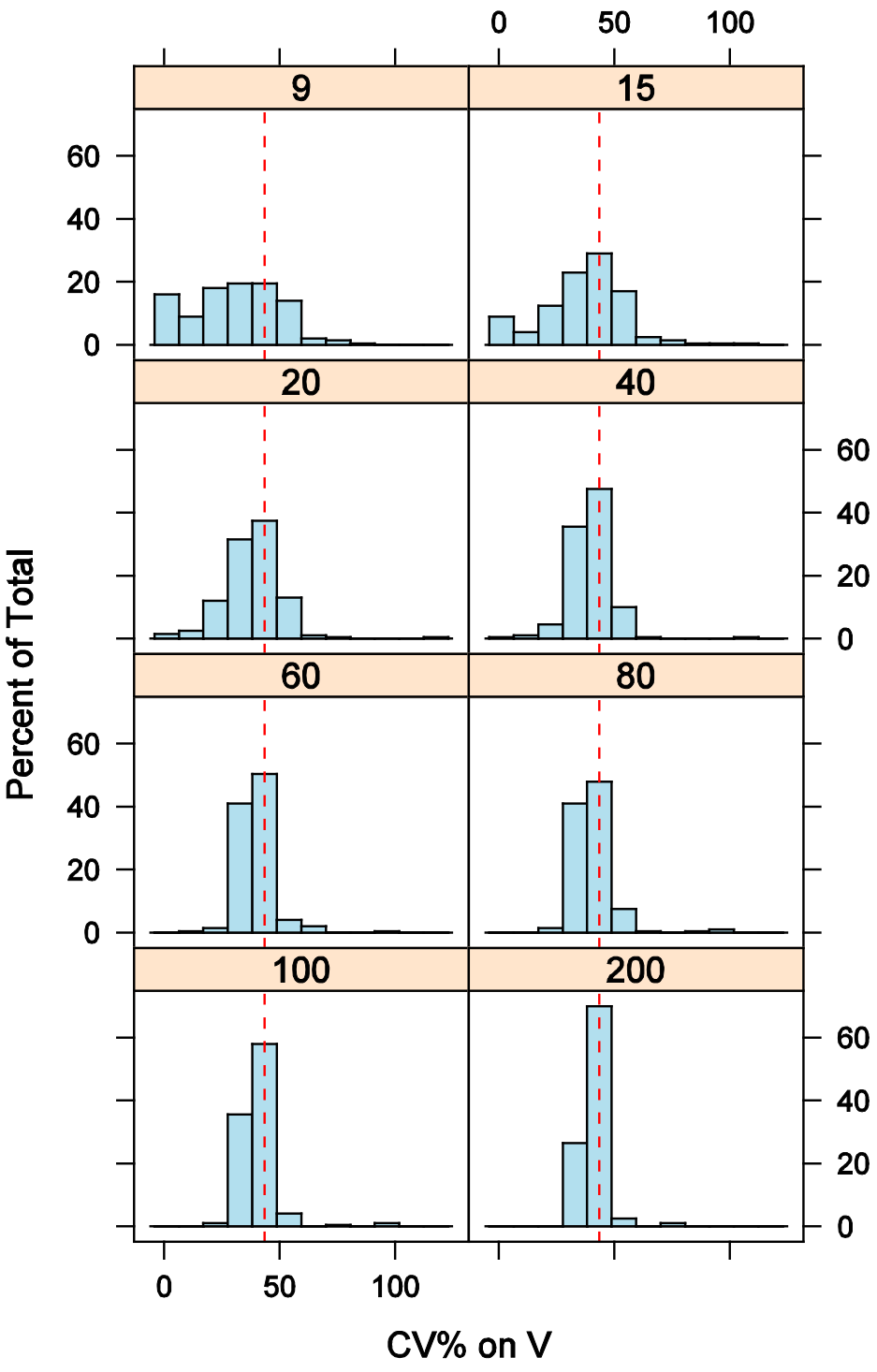


Figure 3-6. Estimated BSV on population volume of distribution in terms of CV% vs. numbers of subjects

The red dash line indicates the CV% = 43.5%.

Table 3-4. RMSE of parameter estimates with various sample sizes

Sample Size	θ_1	θ_2	θ_3	ω_{CL}	ω_V	σ
9	0.00293	0.00020	0.00022	0.01849	0.13761	0.96455
15	0.00263	0.00016	0.00018	0.01836	0.15037	0.74605
20	0.00215	0.00014	0.00016	0.01227	0.12270	0.65776
40	0.00143	0.00009	0.00011	0.01084	0.10307	0.84215
60	0.00128	0.00008	0.00010	0.00637	0.08464	0.47897
80	0.00123	0.00007	0.00010	0.00740	0.09251	0.38901
100	0.00102	0.00006	0.00009	0.00775	0.09318	0.40640
200	0.00082	0.00005	0.00005	0.00427	0.05675	0.35409

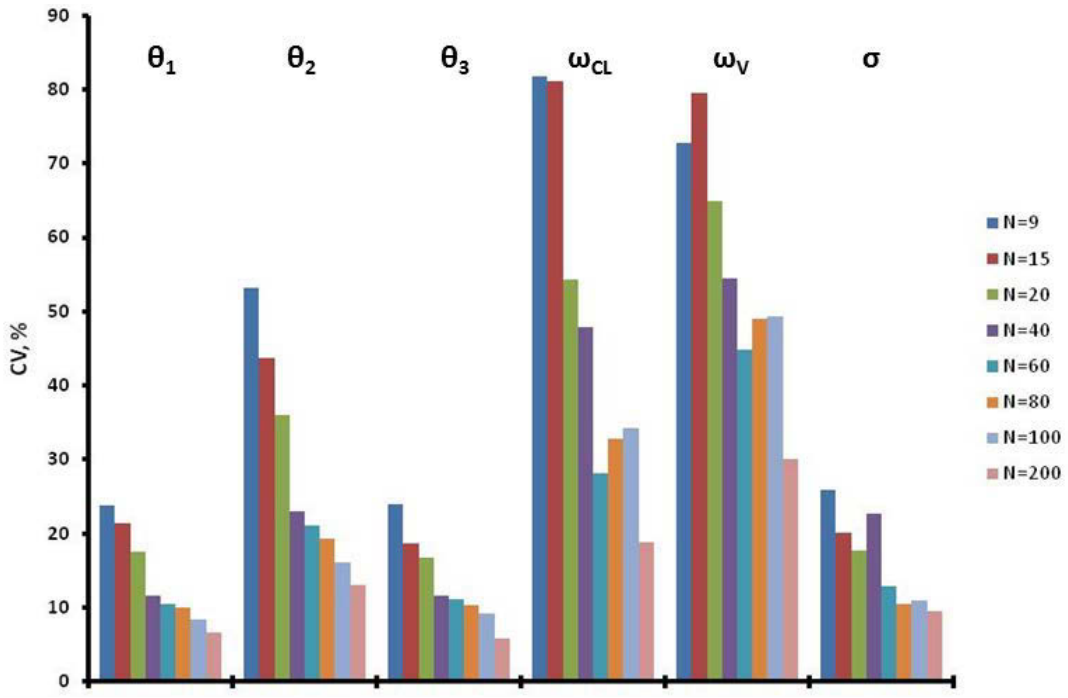


Figure 3-7. Precision of parameter estimates in different sample size groups

Precision is expressed in terms of CV%.

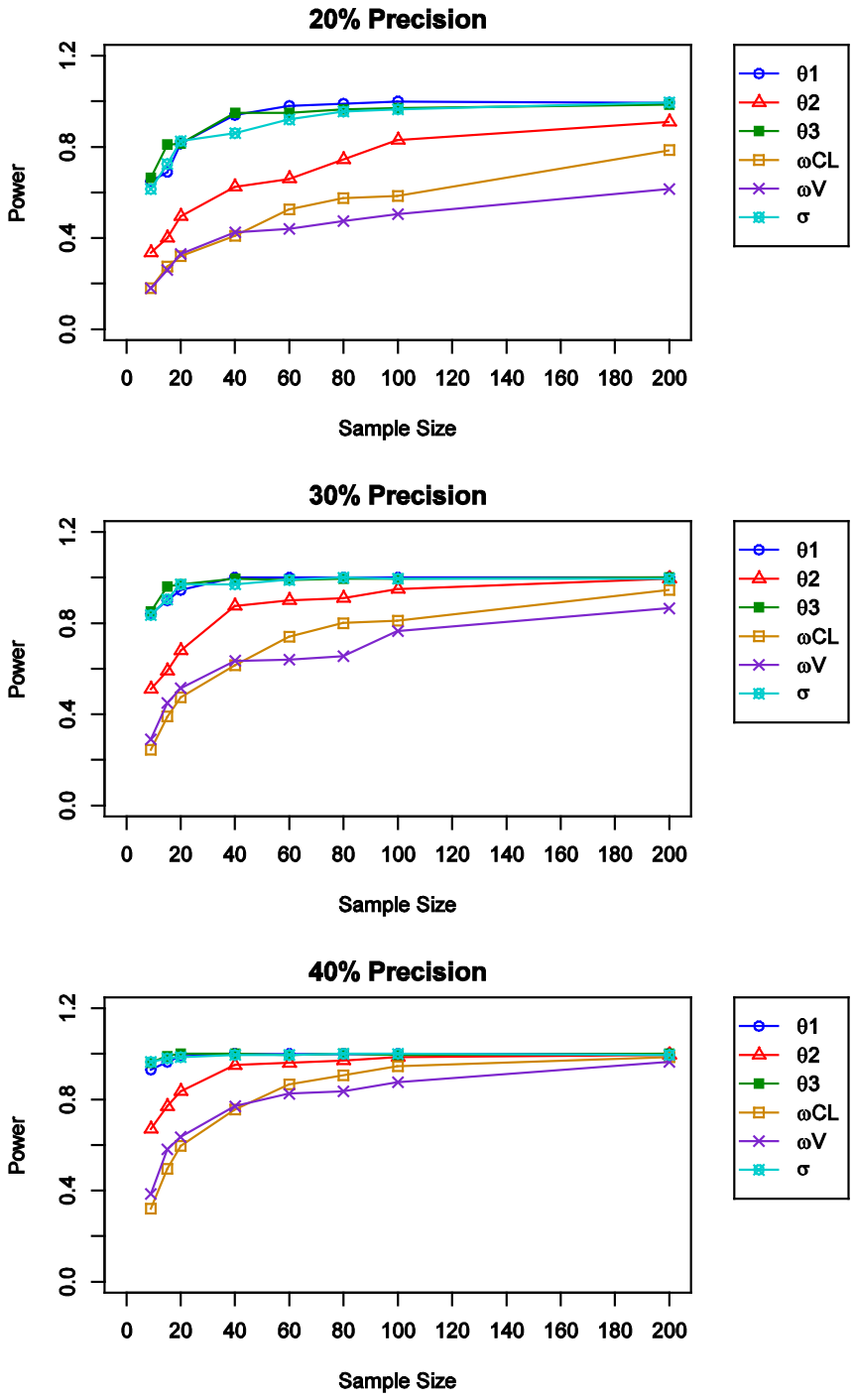


Figure 3-8. Sample size vs. success rate in parameter estimation ("power") at different precision levels of 20%, 30% and 40%

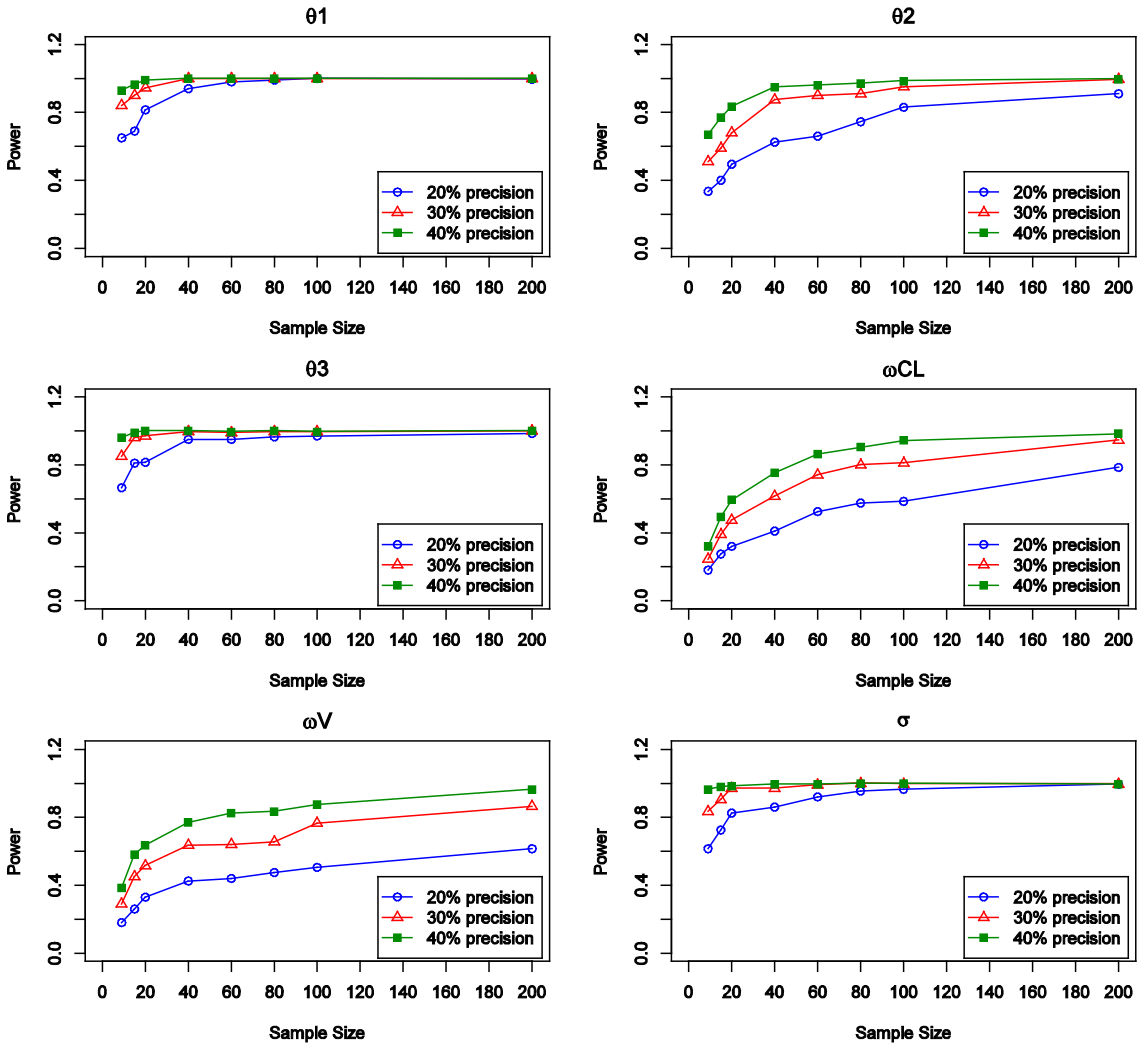


Figure 3-9. Sample size vs. success rate ("power") at different precision levels of 20%, 30% and 40% for each estimated parameter in the model

Covariate effect determination

The power to detect a covariate effect was determined by comparisons of objective function values for the analyses of 200 independent datasets. Objective function value comparisons were made between the structural base model and the covariate model under different sample sizes and three different significance levels. **Figure 3-10** shows the power to detect three covariate effects: body weight on CL, postnatal age (PNA) on CL, and body weight on volume of distribution. Being the most important size covariate for pharmacokinetic parameters in pediatric population, weight can be detected with a higher power using a relatively small sample size. To detect the effect of weight on CL with power of 0.8, 9 subjects would be needed at the level P = 0.05; and 15 subjects would be sufficient at a more stringent level of P = 0.001. The sample size (with significance level) required for the detection of all three covariate effects with power > 0.8 were 20 (P = 0.05), 40 (P = 0.01) and 60 (P = 0.001), respectively.

Discussion

A well determined sample size for a clinical study based on a specified design is considered ethical by limiting the required number of patients while increasing the study power. PopPK studies have the potential to fail by providing unreliable results due to an inadequate study design and a low statistical power. To include the right number of patients is key to solving this problem. To date, most traditional approaches to sample size determination are based on hypothesis testing: A certain degree of difference in PK parameter estimations must be specified beforehand, and then the sample size needed to detect this difference between two or more subgroups is calculated based on Type I and Type II error levels [135-137]. The impact of sample size on the detection of the relationship between covariates and PK parameters has also been studied. Potential covariates were investigated only as categorical factors, such as gender, race, binomial response, age brackets or other category of variable [123, 126]. For PopPK studies in pediatric populations, the concerns are usually regarding the precision and bias of the obtained estimates, instead of a hypothesis or the statistical differences between the parameter estimations. Also, the impact of age and size as continuous covariates on the PK parameters is of great interest. Therefore, we believe the full model-based simulation approach presented here is a more appropriate method for sample size estimation in PopPK studies.

Population pharmacokinetic studies in premature infants have special aspects that must be taken into account for a rational and ethical design. In young pediatric patients, the number of blood samples from each patient and sampling times usually cannot be controlled and predetermined due to therapeutic constraints under which the study is performed. Taking advantage of an aliquot from leftover from blood specimens drawn for therapeutic purposes may be a new way to generate pharmacokinetic information while minimizing patient risk. This sampling strategy determines that the number, and allocations of blood samples collected from each patient, will not be identical. The population modeling process forms the basis for analyzing sparse and unbalanced data

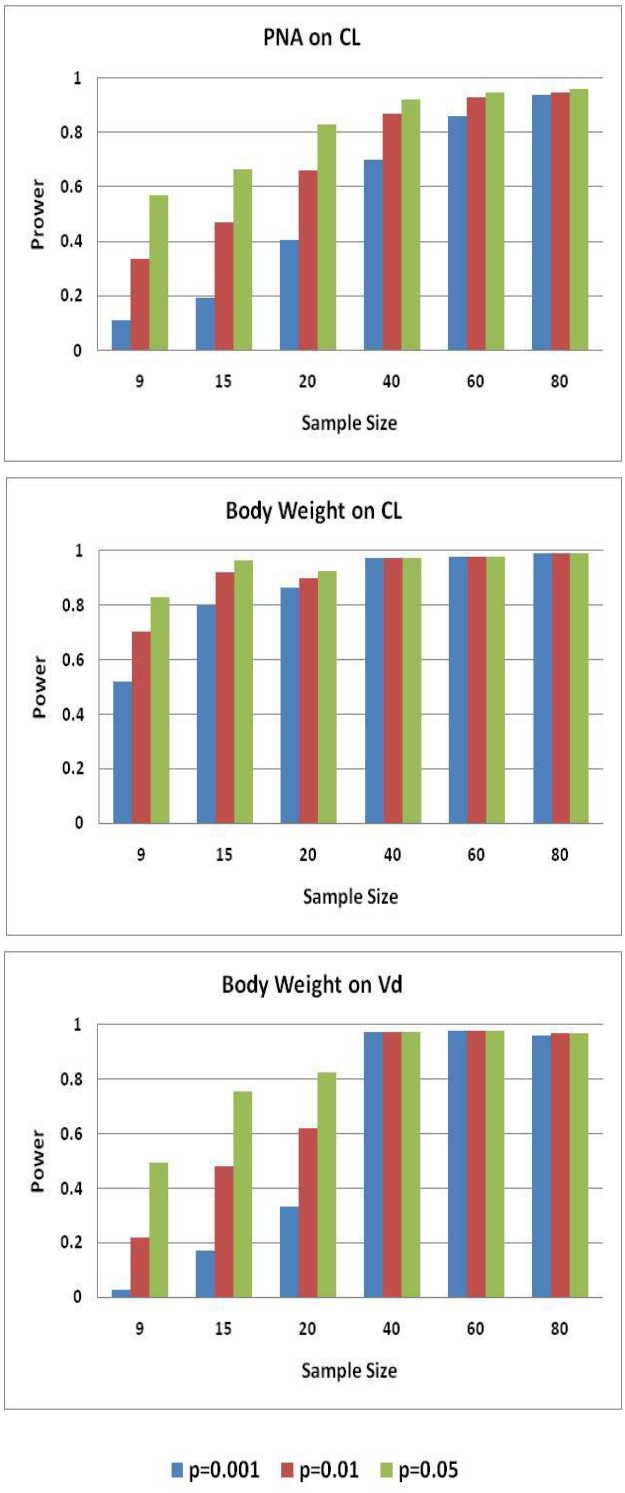


Figure 3-10. Sample size effect on power to detect covariate effects

from multiple patients. Although there has been no such experience on sample size estimations based on an opportunistic blood sampling design in premature neonates group, the model-based simulation approach allows us to explore the impact of sample size on the performance of a PopPK study with fixed model properties. We can optimize the experimental process by comparing and evaluating the predicted outcomes from various designs that cannot be practically explored through clinical testing, thereby facilitating decision making [68].

One potential limitation of clinical trial simulations is the reliability of the prior information included in the analysis. In the current study, the set of covariates generated for each simulated individual should adequately reflect the "real world." Size and age are two important time-varying covariates in neonatal studies that play significant roles in the prediction of individual pharmacokinetic behavior (CL and Vd). Unlike adult studies, multiple-dose studies with neonates cannot neglect characterizing the infants' changes in size even for a short time. In addition, another fixed effect, dose, is also based on the change in body weight. Due to the lack of a demographic database for research in premature infants with very low birth weight, creating datasets with physiologically reasonable covariates was the first challenge. To account for the difference of growth velocity in body weight, all subjects were stratified by 100 g birth weight intervals. Prior knowledge of longitudinal growth of very low birth weight infants was included in the covariate model [132]. A reference growth chart was derived from 1,660 premature infants whose birth weight and gestational age showed similar ranges to in our study. Throughout the simulation procedures, mean parameter estimates for covariate effects with our generated datasets were in strong agreement with the true values reported by the publication used as reference for our PopPK model, suggesting that we successfully created a physiologically reasonable virtual patient population of very premature infants. However, from the available growth charts, anthropometric measurements for birth weights under 500 g were not available. We therefore used the growth velocity for the smallest available birth weight interval of 501 to 600 g for all individuals with a birth weight < 600 g. Another neglected aspect in our simulation was the fact that the body weight simulation did not describe the physiological weight loss commonly observed in newborns during the first week after birth. However, the average daily weight gain partially, if not fully, accounted for this temporary weight loss. Moreover, through the uniform distributed simulation, ~75% of the individuals were older than one week of PNA and had already overcome the dip in weight. No significant effect was expected in the results; though the assumptions made here might lead to a later bias in variance component estimation, they were not a determining factor.

In this simulation study we examined 8 sample sizes from 9 to 200 with 200 independent datasets for each of them. Compared to existing methodologies used in sample size determination for population pharmacokinetic studies, our study displays some unique features. First, a mixed and unbalanced sampling scheme was proposed, which means both sparse and intense sampling existed; the sampling time and the number of samples from each subject were variable. This feature is close to the expected reality of specimen sampling in our future studies in premature infants. Our study evaluated the unusual sampling design through a series of simulations and guided the selection of an

appropriate sample size based on an acceptable success rate of estimation in the sense of "close enough" parameter estimation (similar to the traditional "power" concept but not identical). Moreover, the current simulation study investigated power for continuous covariates as predictors of pharmacokinetic parameters as a function of sample size. Additionally, we performed more complex simulations to mimic the pediatric clinical setting in the analysis, with time-variant covariates and body weight normalized dose levels.

The simulation was fully model-based and took advantage of the estimation methods in NONMEM VI. The fixed-effect PopPK parameters were generally well estimated in terms of accuracy and precision across all tested sample sizes while estimations in variance parameters had larger variability and were more often "off target" compared to those in fixed effects. This finding was fairly consistent with results from other studies [120, 127, 138]. The BSV estimates were all negatively biased except for one value—ω_{CL} when sample size = 100 (See **Table 3-3** and **Figure 3-4**). The possible reason is that all simulations were based on a model which was a simplified form of the real system based on some assumptions. Failing to account for any variability in the model may lead to considerable bias in estimations of variance components, especially for the between-subject variability [136]. In this study, the magnitude of variability employed in the concentration simulation process was derived from the true patients used for original model development; however, we used another covariate model to generate key covariates, which conserved the correlation between PNA and body weight during the 14 days follow-up. These covariates were then used as input in the simulations of longitudinal data and parameter re-estimations. Although the generated datasets adequately characterized the major aspects of the real patient population, they do not necessarily represent atypical patients and the large variability of the "real patients". This covariate model partially contributed to the inaccurate estimations of BSV parameters in our study.

Considering the performance in parameter estimation and covariate effect detection, the minimum sample size required for a theophylline study in premature infants was determined by the desired precision of parameters of interest under a given blood sampling design. If the bias in parameter estimations in terms of %MPE was set as a cut-off criterion no greater than 15%, and ≤ 25% and ≤ 50% were accepted measures of precision for fixed effect and variance parameters, respectively, a sample size of 40 subjects was sufficient. At a sample size of 40 subjects, the power to detect the covariate effect was > 80% at a significance level of P = 0.01.

It has been widely recognized that the sample size estimation of a PopPK study is considerably influenced by changing study design factors and model properties, such as the allocation of blood sampling times, number of blood samplings, estimation algorithms in NONMEM, and the magnitude of between-subject variability and residual variability [124, 127, 139, 140]. For example, FO and FOCE methods are two commonly used NLME estimation algorithms to obtain parameter estimations in NONMEM VI. FO might result in considerable bias in parameter estimation due to approximation of the true likelihood function. Compared to FO, FOCE is considered to perform better and with less

bias in parameter estimation when there is large between-subject and residual variability. Therefore, FOCE was selected as the only estimation algorithm throughout the simulation study. To focus on our target question of interest, the impact of other estimation methods in NONMEM was not investigated in the current simulation study. However, the proposed approach also can be used to assess the performance of other NLME estimation methods, such as first order conditional estimation with interaction (FOCEI) in NONMEM VI, stochastic approximation expectation maximization (SAEM), Monte Carlo importance sampling (IMP), Monte Carlo importance sampling assisted by mode a posteriori (IMPMAP), and Markov chain Monte Carlo Bayesian (BAYES) in NONMEM VII. The application of D-optimality-based, limited sampling schemes is well acknowledged for increasing trial efficiency and minimizing the necessary number of blood samples by providing informative sampling designs [141-143], which obviously is beyond the discussion of this chapter due to our specified opportunistic sampling characteristics. However, we can still expect a reduction in the number of patients required for reaching the same level of success rate/"power" using an extensive blood sampling design or optimal sampling scheme compared to the sparse and opportunistic sampling design. The proposed approach may also prove valuable in studying other drugs of interest in premature infants if appropriate prior knowledge is available.

Conclusions

In conclusion, we developed a full model-based simulation approach to the sample size determination for PopPK studies in premature neonates. A mixed and unbalanced sampling design was used in the analysis. For a desired accuracy, precision and study power, the appropriate number of patients with a specified sampling design was determined using the proposed approach. While the accuracy and precision of parameter estimation were shown to benefit from increases in the number of subjects in the evaluated observational study approach, designs with 20 premature neonates were shown to be inadequately powered to allow for accurate and precise estimation of PopPK parameters. Designs with > 40 subjects were required for $\leq 15\%$ in bias, and $\leq 50\%$ in precision for parameter estimations of both fixed and random effects with adequate power. This result will be useful in selecting samples sizes for upcoming clinical studies in premature neonates.

CHAPTER 4. POPULATION PHARMACOKINETIC ANALYSIS OF CAFFEINE IN PREMATURE NEONATES WITH APNEA

Introduction

Apnea of prematurity (AOP)

Apnea of prematurity (AOP) is one of the major concerns in premature neonates. AOP is defined as the cessation of breathing that lasts for 15 or 20 seconds and is usually accompanied by dangerous hypoxia and/or bradycardia [144]. The incidence of AOP increases with increasing prematurity of birth. It affects approximately 85% of neonates with birth age < 34 weeks gestational age (GA), while reaching nearly 100% in infants born at < 29 weeks GA or having birth weight < 1 kg [145]. AOP is commonly treated with methylxanthines such as caffeine (1, 3, 7-trimethylxanthine) and theophylline (1, 3-dimethylxanthine) as respiratory stimulants.

Application of caffeine in patients with apnea of prematurity

Being the current first-line pharmacotherapy for the treatment of AOP, caffeine has been used frequently for more than three decades [99]. Caffeine therapy is believed to reduce the number of apnoeic episodes, the duration of respiratory support of continuous positive airway pressure (CPAP), the incidence of bronchopulmonary dysplasia (BPD) and, from a long-term view, the morbidity of neurodevelopmental disability in very low birth weight premature infants [144-146]. In most cases, its administration to neonates and infants was empirical and off-label, the first and to date the only commercially available caffeine product, CAFCIT®, was approved by the U.S. Food and Drug Administration (FDA) at the end of the last century. This medication is labeled for the short-term treatment of AOP in premature infants 28-33 weeks GA. Therefore, dosing guidance remains empiric and variable for those extremely low birth weight (ELBW) infants (birth weight < 1000 g), which are also the youngest premature neonates (23 < GA < 28 weeks) [147]. The percentage of infants who are ELBW and very low birth weight (< 1500 g) has steadily been increasing in the last ten years. A new cohort of neonates, so-called fetal infants, whose birth weight is < 500 g, is also growing. This increase might be associated with multiple reasons, such as:

- Increased incidence of very premature birth is associated with multiple pregnancies and multiple births caused by assisted reproductive technology (ART);
- Improved neonatal survival due to technological advancement in perinatal and neonatal care; and
- Progress in the medical management, increases in early cesarean section and

induction of labor due to pregnancy complications or health problems.

Because of the increasing survival of ELBW infants, the burden of morbidity from AOP and its associated BPD is growing. However, there is only a limited amount of pharmacokinetic (PK) data for caffeine available for this newly emerging population of premature infants. This population is susceptible to developmental changes that can affect the disposition of drugs. Many neonatal conditions in physiology and pathology are unique, and their pharmacokinetic characteristics have rarely been investigated. As discussed in Chapter 1, linear extrapolation of dosing regimens based on body weight or body surface area may not be appropriate and is frequently associated with adverse events and lack of efficacy. To develop a sound, scientifically-based caffeine pharmacotherapy in premature infants is therefore urgently needed for this patient population to ensure safe and effective treatment of AOP. A preferable approach for pharmacokinetic studies in pediatrics is the population pharmacokinetic modeling approach [73]. The PopPK approach allows for utilizing sparse and unbalanced data collected during routine clinical care, including therapeutic drug monitoring (TDM), from individual patients to determine factors that may influence the drug behavior in the human body. Therefore, using the PopPK method can help to overcome the scientific, logistic and ethical limitations of traditional PK studies in premature neonates.

Objective

A number of clinical studies using caffeine have been performed in patients with AOP. Caffeine is generally considered a safe and effective medication in the NICU. However, the optimal dose is still not known. Future studies will focus on maximizing therapeutic benefits while minimizing toxicity through a series of dose selection analyses [101, 148]. The objectives of the current study were: (1) to develop a PopPK model of caffeine in premature neonates, (2) to determine the typical PopPK parameters and associated between-subject variability of caffeine, (3) to assess and identify potential sources of variability of PK behavior for caffeine among premature neonatal patients throughout infancy and (4) to use this PopPK model to further facilitate the development of optimal dosing regimens through simulation—particularly, to correlate steady state concentrations with response at different dosing regimens for various age/weight groups.

Methods

Approval

The current study received approval by the Eastern Virginia Medical School Institutional Review Board, Norfolk, VA.

Study design and patient population

A total of 560 caffeine concentration measurements were gathered from 88 hospitalized patients with the main diagnosis of apnea of prematurity from July 2008 to December 2008 in the Neonatal Intensive Care Unit of Children's Hospital of the King's Daughters, Norfolk, VA. All pharmacokinetic data were obtained retrospectively from the medical records and routine therapeutic drug monitoring. The subjects included in the PopPK analysis received repetitive intravenous (IV) and/or oral administration of caffeine. Caffeine was given as an initial IV loading dose over 30 minutes, followed by maintenance doses every 12 or 24 hour via 10 minutes IV infusion or orogastric administration. Loading doses ranged from 6.5 to 9.5 mg/kg, while maintenance doses ranged from 3.1 to 28.6 mg/kg/day. All blood PK samples were drawn before the morning drug administration. All patients had reliable dosing and sampling collection date and time information recorded and at least one associated measurable concentration of caffeine. The covariates collected for each patient included birth weight (BW), body weight (WT), GA, postnatal age (PNA), postconceptional age (PCA), gender, race and the use of respiratory support. Demographic characteristics and relevant clinical profiles of pharmacokinetic data were collected and precisely documented.

Assay methodology

Plasma concentrations of caffeine were determined in the clinical laboratory at Children's Hospital of the King's Daughters with a homogenous spectro-photometric method using a Syva enzyme-multiplied immunoassay technique performed on an ARCHITECT® c8000™ analyzer by Abbott Diagnostics (Dallas, Texas). The assay range was from 1.0 mg/L to 30.0 mg/L, while a dilution would be performed on the PK samples with a concentration > 30.0 mg/L. The inter-assay coefficient of variation was 4.2% at a concentration level of 11.0 mg/L.

Population pharmacokinetic analysis

Computer and software

The population pharmacokinetic analysis of caffeine was carried out by nonlinear mixed-effects modeling using a NMQual 6.4.1 (Metrum Institute, Augusta, Maine) installation of NONMEM version VI, Level 2.0 (ICON Development Solutions, Ellicott City, Maryland) with a GNU Fortran 77 (g77) version 2.95. Data summary and figures were prepared with Xpose [149], Census Version 1.0 [150], R 2.8.1 (http://www.r-project.org), or SAS 9.1.

Structure model development

The first order conditional estimation (FOCE) method within NONMEM was used for the estimation in the PopPK analysis. A one-compartment model based on the

available literature report was used to describe the caffeine concentration-time profile [151-153]. The structural model to be tested was selected as a one-compartment model with first-order absorption and first-order elimination without lag time. Since the concentration data were collected during routine drug monitoring and consisted of trough data, no information from the absorption phase after oral administration was available for the evaluation of absorption rate constant (ka); therefore, a fixed value of 10 hr^{-1} was employed to represent its rapid absorption suggested in the literature[152-154]. Caffeine concentration data were log transformed prior to the PopPK analysis. Between-subject variability on the pharmacokinetic parameters of caffeine was assumed to follow a log-normal distribution and was modeled with an exponential error model.

$$P = \theta \times (e)^{\eta}$$

$$\eta \sim N(0, \omega^2)$$

Where θ represents the typical value of a population pharmacokinetic parameter, P is the true but unknown value of θ in the individual. η represents the population parameter variability in model parameters, independently and randomly distributed with mean zero and variance ω^2. Stepwisely, between-subject variability on all PK parameters was added or removed from the model. Residual error was initially modeled with an additive error model on the log-transformed concentration data.

$$\ln(C_{obs,ij}) = \ln(C_{pred,ij}) + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

Where $C_{obs,ij}$ is the i^{th} observed concentration in the j^{th} subject, $C_{pred,ij}$ is the i^{th} model predict concentration in the j^{th} subject and ε_{ij} is the deviation of $\ln(C_{obs,ij})$ from $\ln(C_{pred,ij}. \varepsilon_{ij})$. ε is a normally distributed random variable with an average value of 0 and variance of σ^2.

A diagonal covariance matrix for the between-subject variability was initially used. After the initial diagonal covariance matrix was identified for the base model, off-diagonal correlations were also tested if a scatter plot correlation matrix indicated significant correlation between individual parameters calculated by the posterior conditional estimation (POSTHOC) technique within NONMEM. Once the between-subject variability covariance matrix was determined, the residual error model was also further evaluated with a proportional and additive error model on the log-transformed concentration data.

Initially, the basic model was evaluated without any covariates. Due to the fact that demographic factors such as weight and age play a significant role in determining pediatric pharmacokinetic parameter estimates and exploratory modeling results, body weight was included *a priori* during the base model development, along with fixed allometric exponents of 0.75 and 1 for clearance (CL) and volume of distribution (V), respectively. Fixing one of the covariate-parameter relationships to an allometric

expression allows estimation of effects of other covariates if these covariates are highly collinear [73]. The use of these coefficients is supported by fractal geometric concepts and observations from diverse areas in biology [39, 73].

$$CL = \theta_{CL} \times \left(\frac{WT}{Median}\right)^{0.75}$$

$$V = \theta_{V} \times \left(\frac{WT}{Median}\right)^{1}$$

Covariate model development

Once the base model was identified, the influence of subject-specific covariates on the estimated PK parameters was evaluated. The covariates screened included PNA, GA, PCA, BW, low gestational factor (LGA), gender and race. Prior to the covariate model development, a scatter plot correlation matrix was developed to identify any high intercorrelation among covariates.

For continuous covariates, scatter plots of individual PK parameter estimates against covariates overlaid with a LOESS smooth line were used to help identify functional relationships. For categorical covariates, box and whisker plots of individual PK parameters for each of the groups were used to identify differences between groups. Continuous covariates were modeled using proportional or linear relationships in a median-centered manner:

$$P = \theta_1 + \theta_2 \times (COV - Median)$$

$$P = \theta_1 + \theta_1 \times \theta_2 \times (COV - Median)$$

Where θ_1 represents the typical value of a PK parameter in an individual with the median value for the covariate (COV) and θ_2 represents the coefficient for the relationship with the covariate. P is the individual pharmacokinetic parameter. If the scatter plot between the covariate and the individual PK parameter indicated a log-linear or exponential relationship, the following power model was used:

$$P = \theta_1 \times \left(\frac{COV}{Median}\right)^{\theta_2}$$

$$P = \theta_1 + \left(\frac{COV}{Median}\right)^{\theta_2}$$

Combinations of proportional, linear and power models were developed as needed. Categorical covariates were modeled using a fractional change model:

$$P = \theta_1 \times (1 + \theta_2 \times COV)$$

Where COV has either the value of 0 or 1.

The covariate model was developed using a stepwise forward addition and backward elimination approach. First, covariates were added to the base model incrementally in a univariate fashion and were tested to determine whether there was a statistically significant decrease in objective function value of 3.84 ($p < 0.05$) based on the Chi-square test with one degree of freedom. Covariates that demonstrated significant PopPK model improvement were considered for the next iteration of covariate model development. The covariate model demonstrating the greatest improvement in the PopPK model was incorporated into the base PopPK model while remaining covariates were re-evaluated incrementally. This process was repeated until none of the remaining covariates provided significant improvement to the PopPK model.

Following determination of the fully parameterized PopPK model, a backward elimination approach was used to evaluate if all covariates included in the full model continued to provide significant influence on PK parameter estimations. Thus the included covariates were sequentially removed from the full model to determine if there was significant model deterioration. A more stringent p-value of 0.01, based on the Chi-square test with one degree of freedom, was used during the backward elimination process to avoid false-positives. This process was repeated until the model contained the minimum number of parameters that produced no significant PopPK model deterioration.

Covariate model development also was guided by considering physiological and pharmacological mechanisms, reduction in the between-subject variability on the corresponding PK parameters and improvement of goodness-of-fit plots.

Model selection

The following criteria were applied during model development to identify an improved model:

- A significant reduction in the objective function value based on the likelihood ratio test. A decrease of > 3.84 points in objective function value was considered significant ($p < 0.05$) for addition of one model parameter.

- The improvement in pharmacokinetic parameter estimation, such as a decrease in the estimated standard error for model parameters, a decrease in the magnitude of the between-subject variability for pharmacokinetic parameters and/or a decrease in the magnitude of the residual error.

- Evaluation of the goodness-of-fit plots, including a less systematic or narrower distribution of individual predicted versus observed dependent variable and a random distribution in the residuals/weighted residuals versus the predicted dependent variable, versus time or versus covariates.

Population pharmacokinetic model qualification

Bootstrap analysis

A nonparametric bootstrap analysis was performed to evaluate the model performance internally. First, 500 bootstrap data files were created from the original NONMEM data file with repeatedly random sampling with replacement, consisting of the same patient sample size as the original NONMEM data file. Then the population parameters were estimated for each of the 500 bootstrap data files using the final covariate model. The same optimization method was used in the estimation as that in the final model. Based on these estimations, the median and 95% confidence intervals (CIs) for each parameter were derived from the successfully converged bootstrap runs using the percentile method and compared with the estimates of the original index dataset.

Visual predictive check

The predictive performance of the final covariate model was assessed by conducting a visual predictive check. Using the final PopPK model with the parameter estimates and their distributions, 500 simulations were performed with the original dataset to preserve the covariate vector. The 90% prediction confidence intervals (5% and 95% percentiles) were determined based on the simulated plasma concentrations data. Since actual times were used for analysis, intervals were binned around the most frequent times after dose as determined by the data structure of the original dataset. Prediction confidence intervals were plotted and overlaid with the observed concentrations. If approximately 90% of the observed values judged by visual inspection laid within the 90% confidence intervals for the simulated concentrations with comparable spread, then the PopPK model was considered to have strong predictive value.

Posterior predictive check

As a posterior predictive check, a separate Monte Carlo simulation was performed in NONMEM with six representative subjects chosen from the model building dataset. Dosing regimens used in this simulation were selected based on a previous dose–optimization analysis. The results were then evaluated and compared to the observed caffeine levels.

Shrinkage

The shrinkage for the individually estimated concentrations (epsilon shrinkage) and parameters (eta shrinkage) were calculated [155]. Epsilon shrinkage was calculated as 1 – SD (IWRES), where IWRES is the individual weighted residual. Eta shrinkage was calculated as 1 – [SD (EBEs)]/ω, where EBEs stand for empirical Bayes estimates (or individual POSTHOC parameters) and ω is the population model estimate of the SD in eta. Although no formal criterion exists, a level of < 0.3 was used to conclude "no relevant shrinkage."

Dose-optimization study

The final covariate PopPK model for caffeine (and its parameter estimates) was subsequently applied to a dose-optimization simulation analysis. Various dosing levels with 12 or 24 hour dosing intervals were evaluated through Monte Carlo simulations using Trial Simulator Version 2.2 (Pharsight Corporation, Mountain View, CA).

In a meta-design, all subjects were divided into 11 sub-groups based on their body weight and PCA. Twelve treat arms at 2, 2.5, 3, 4, 5, 6, 7, 8, 9, 10, 12 or 14 mg/kg/day caffeine base with dosing intervals of 12 or 24 hours given for 15 days were selected in the evaluation based on a preliminary simulation analysis. The covariate distributions for the simulated PCA and body weight were consistent with those in the observed caffeine data set in terms of mean, standard deviation and range. A joint distribution between PCA and WT for each sub-group was employed in the covariate model. Correlation coefficients directly derived from the caffeine data set were used in the model building to account for the slight difference in collinearity between PCA and body weight in each sub-group. For each evaluated scenario, 400 replicates were generated. Then the peak and trough concentrations on the 15^{th} day for each dosing regimen were summarized and compared to the therapeutic target concentrations of caffeine. The probabilities of achieving the desired therapeutic target were estimated for each group.

Results

Pharmacokinetic data

The final caffeine analysis dataset consisted of a total of 560 PK samples collected from 88 subjects with a GA of 23–31 weeks and a diagnosis of apnea of prematurity. The dosing and baseline demographics of the subjects are summarized in **Table 4-1**. The duration of caffeine treatment ranged from 9 to 107 days, with a median of 48 days. The median number of concentration measurement provided is 6 per subject. The majority of patients (70 out of 88 subjects) contributed 4-10 PK observations, while 13 subjects had 1-3 PK observations and 5 subjects had 11-17 PK observations. It should be noted that only 3 subjects had a single PK observation (**Figure 4-1**).

Structural model

One-compartment models with very fast first-order (ka = 10 h^{-1}) absorption, and first-order elimination were tested using the subroutines ADVAN2 and TRANS2 within NONMEM. Various combinations of between-subject variability (exponential) on CL, V and F1 with or without off-diagonal correlation were explored. The residual error term was initially modeled as an additive error on the log-transformed concentrations. The proportional and additive residual error model was also tested once the between-subject

Table 4-1. Subject dosing and baseline demographics summary

Characteristics	N	Median	Range
Gender (male/female)	38/50		
Race (Caucasian/Black/Other)	28/50/10		
Caffeine dose (mg/kg/day)		7.7	3.1-28.6
Duration of caffeine therapy (day)		48	9-107
Caffeine concentration (mg/L)	560	24.4*	7.9-42.7
Concentration measurements per patient		6	1-17
PNA (day)	88	39	1.0-116
GA (week)	88	26	23-31
PCA (week)	88	32	24-42
BW (kg)	88	0.84	0.38-1.7
WT (kg)	88	1.3	0.36-3.0

N = number of subjects.
* Presented as mean.

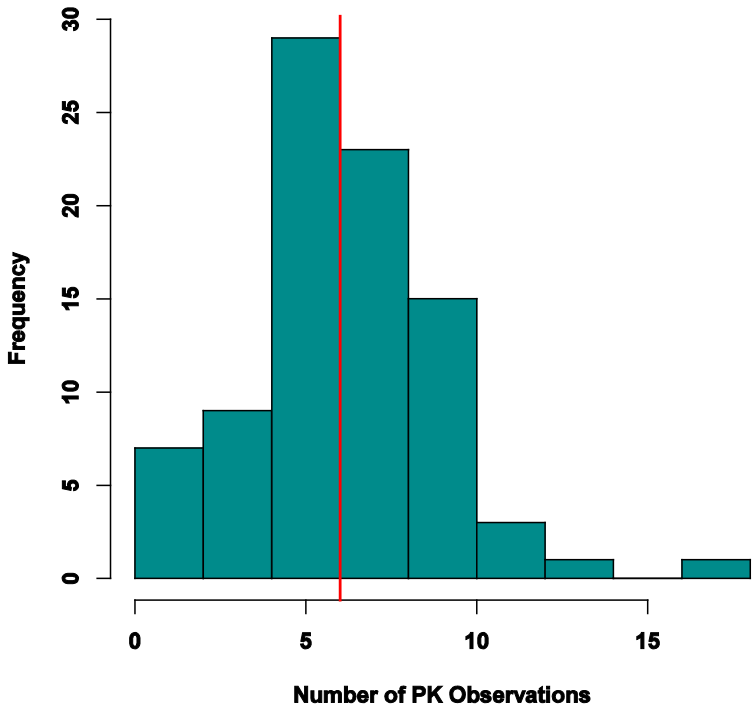


Figure 4-1. Histogram of frequences of number of caffeine concentrations contributed per subject

Red vertical line indicates the median of the number of PK samples.

variability had been selected. The structural base model was finally determined as a one compartment model with no lag-time, first-order absorption, first-order elimination and between-subject variability expressed as exponential terms on CL, V and oral bioavailability (F1). Residual error was modeled as an additive error on the log-transformed concentrations. The models were evaluated based upon the objective function value, goodness-of-fit plots, parameter estimates, precision of parameter estimates, between-subject variability and residual error.

Size and maturation-related covariates (GA, PNA and PCA) are usually two important aspects associated with pediatric analysis, and collinearity is commonly observed (**Figure 4-2)**. In neonatal studies, both covariates have profound but indistinguishable impact on parameter estimation. The incorporation of allometric size adjustment by weight centered by the median value of 1.5 kg for both CL and V was evaluated by two approaches, *a priori* inclusion with fixed exponents to 0.75 for CL, 1 for V and estimation as part of the covariate model building. Both approaches gave significant reductions ($P < 0.001$) of the objective function value, 438.5 and 558.8, respectively. The allometric exponential estimates and subsequent parameter estimate of PCA effect on CL are listed in **Table 4-2**.

The pros and cons of these two approaches for size adjustments in PopPK analyses have been well discussed in the literature, particularly in a situation where collinearity exists among the studied covariates [39, 73, 156]. To allow for the evaluation of the influence of other covariates that are collinear with WT on the PK parameters of caffeine, the allometric size adjustment by *a priori* fixing the exponents to 0.75 for weight on CL and 1 for weight on V was finally decided on for use in the base model. The PK parameter estimates of caffeine obtained from the base model are given in **Table 4-3**.

Covariate model

Continuous covariates, including PNA, PCA, GA and BW, were tested in a linear manner or a power function normalized with the median value on CL, V and F1. Categorical covariates including gender and race were tested on CL, V and F1 in a fractional change manner.

In the first forward addition step, PCA on CL (power function), low gestational age (LGA) factor on CL, and LGA on V were found to affect the model fit significantly. PCA on CL showed the greatest reduction (-160.4) in objective function value compared to the base model and was then chosen as the reference model for Step 2. Various LGA factors (when GA = 24, 25, 26, 28, 30 weeks) were evaluated, but only LGA < 25 weeks was found to have a significant effect on both CL and V.

In the second forward addition step, only those covariates that were found significant in the first step were tested. LGA on CL and LGA on V were significant. Among them, LGA on CL produced the greatest decrease in objective function value

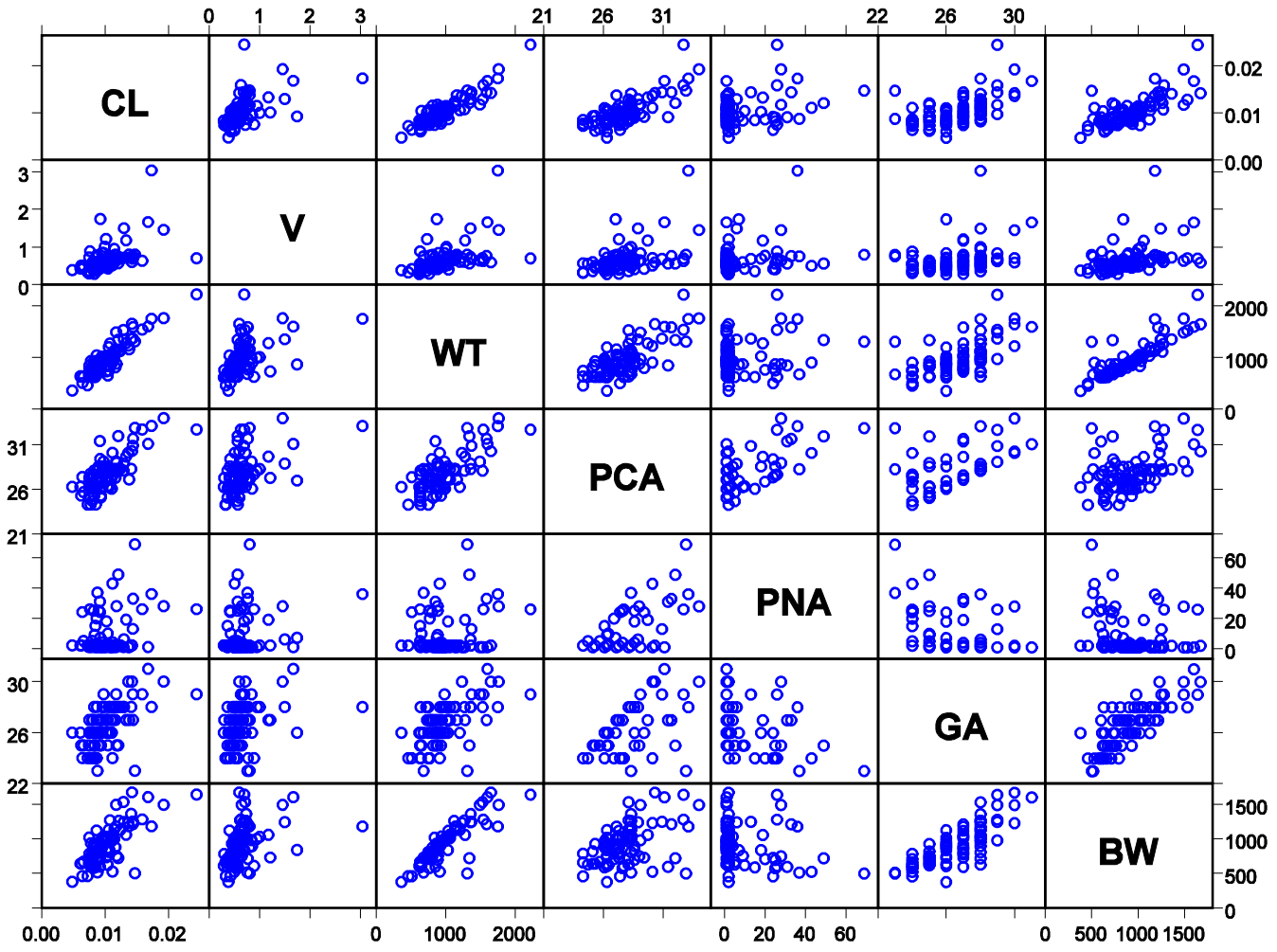


Figure 4-2. Matrix of continuous covariates and PK parameters of the base model

Table 4-2. Population pharmacokinetic parameters of caffeine obtained from the allometric model with estimated or fixed exponentials

Model Parameters	Pre-base Model[a]	Pre-base Model[a] + PCA on CL	Base Model[b]	Base Model[b] + PCA on CL
Clearance (CL, L/hr)	0.0169 (15.3%)	0.0164 (10.4%)	0.0144 (14.9%)	0.0166 (10.3%)
Weight on CL (power)	1.2 (0.057)[c]	0.736[d]	0.75	0.75
PCA on CL	-	1.88[d]	-	1.91 (0.17)[c]
Volume of distribution (V, L)	1.15 (56.5%)	1.19 (52.7%)	0.981 (52.6%)	1.01 (45.8%)
Weight on V (power)	1.45 (0.218)[c]	1.36[d]	1	1
Absolute bioavailability (F1)	0.97 (13.1%)	1.01 (13.5%)	0.824 (14.8%)	1.02 (13.9%)

Parameter estimates are presented as mean (between-subject variability, %).
[a] Exponential of WT on CL and V were estimated.
[b] Exponential of WT on CL and V were fixed.
[c] Presented as mean (standard error).
[d] Standard error not available.

Table 4-3. Population pharmacokinetic parameters of caffeine obtained from the base model

Model Parameters	Parameter Estimate	RSE[a]	BSV[b] (CV%[c])
Clearance (CL, L/hr)	0.0144	2.22%	14.9%
Weight on CL (power)	0.75	NA	NA
Volume of distribution (V, L)	0.981	7.57%	52.6%
Weight on V (power)	1	NA	NA
Absolute bioavailability (F1)	0.824	2.99%	14.8%
Residual error	0.0421	8.53%	20.5%

[a] Relative standard error (RSE), calculated as (Standard Error/Estimate)*100 from NONMEM® results.
[b] Between-subject variability (BSV).
[c] Coefficient of variation (CV%), calculated as $\sqrt{\omega^2} * 100$, where ω^2 is the between-subject variance estimate.

(ΔOFV = -14.3).

In the third step, only those covariates that were found significant in the second step were tested. In the third forward addition step, LGA on V was found significant (ΔOFV = -5.13). Thus this model was then considered the fully parameterized model; it consisted of the statistically significant relationships of a body weight effect on CL (positive power relationship), a PCA effect on CL (positive power relationship), a low GA effect on CL (positive fractional change), a body weight effect on V (linear relationship) and a low GA factor on V (positive fractional change).

Removal of each of these covariates led to a significant deterioration of the model, as indicated by an increase of more than 6.64 points in the OFV compared to the full model. Therefore, each relationship was considered significant. The final model was a one-compartment model with no lag-time, between-subject variability expressed as exponential terms on CL, V and F1, and having the exponential residual error model and covariate effects listed in the previous paragraph.

Additionally, low birth weight factor (LBW) (BW = 0.6, 0.8, 0.9, 1.0, 1.2 kg) and respiratory support (5 rank levels of oxygen support) were also tested as potential covariates but did not appear to be significant. The only covariates found to significantly ($p < 0.01$) affect the PK parameters beyond body weight were PCA on CL and a factor < 25 weeks of gestational age at birth (LGA) on CL and V.

The final models for clearance and volume of distribution are as follows:

$$\text{CL (L/hr)} = 0.0164 * (\frac{WT}{1.5})^{0.75} * \left(\frac{PCA}{32}\right)^{1.96} (* \ 1.18, if \ GA < 25 \ weeks)$$

$$\text{Vd (L)} = 0.94 * (\frac{WT}{1.5})^{1.0} (* 1.57, if \ GA < 25 \ weeks)$$

$$F = 1.0$$

The typical patient in the studied premature neonate population, i.e., a patient with WT of 1.5 kg, PCA of 32 weeks, and GA > 25 weeks, was estimated to have a CL of 0.0164 L/hr and a V of 0.94 L. That is equivalent to 0.0109 L/kg/hr and 0.63 L/kg for CL and V, respectively. The population typical value for absolute bioavailability was estimated to be close to 1.

The PK parameter estimates of caffeine obtained from the final model are given in **Table 4-4**. Selected goodness-of-fit plots are presented in **Figure 4-3**, **Figure 4-4**, **Figure 4-5**, **Figure 4-6**, and **Figure 4-7**.

Table 4-4. Population pharmacokinetic parameters of caffeine obtained from the final model

Model Parameters	Parameter Estimate	RSE[a]	BSV[b] (CV%[c])
Clearance (CL, L/hr)	0.0164	2.29%	8.9%
Weight on CL (power)	0.75	NA	NA
PCA on CL (power)	1.96	8.72%	NA
Low GA on CL (fraction)	1.18	3.46%	NA
Volume of distribution (V, L)	0.94	7.51%	42.3%
Weight on V (power)	1	NA	NA
Low GA on V (fraction)	1.57	16.2%	NA
Absolute bioavailability (F1)	1.0	3.31%	14.2%
Residual error	0.0318	8.84%	17.8%

[a] Relative standard error (RSE), calculated as (Standard Error/Estimate)*100 from NONMEM® results.
[b] Between-subject variability (BSV).
[c] Coefficient of variation (%CV), calculated as $\sqrt{\omega^2} * 100$, where ω^2 is the between-subject variance estimate.

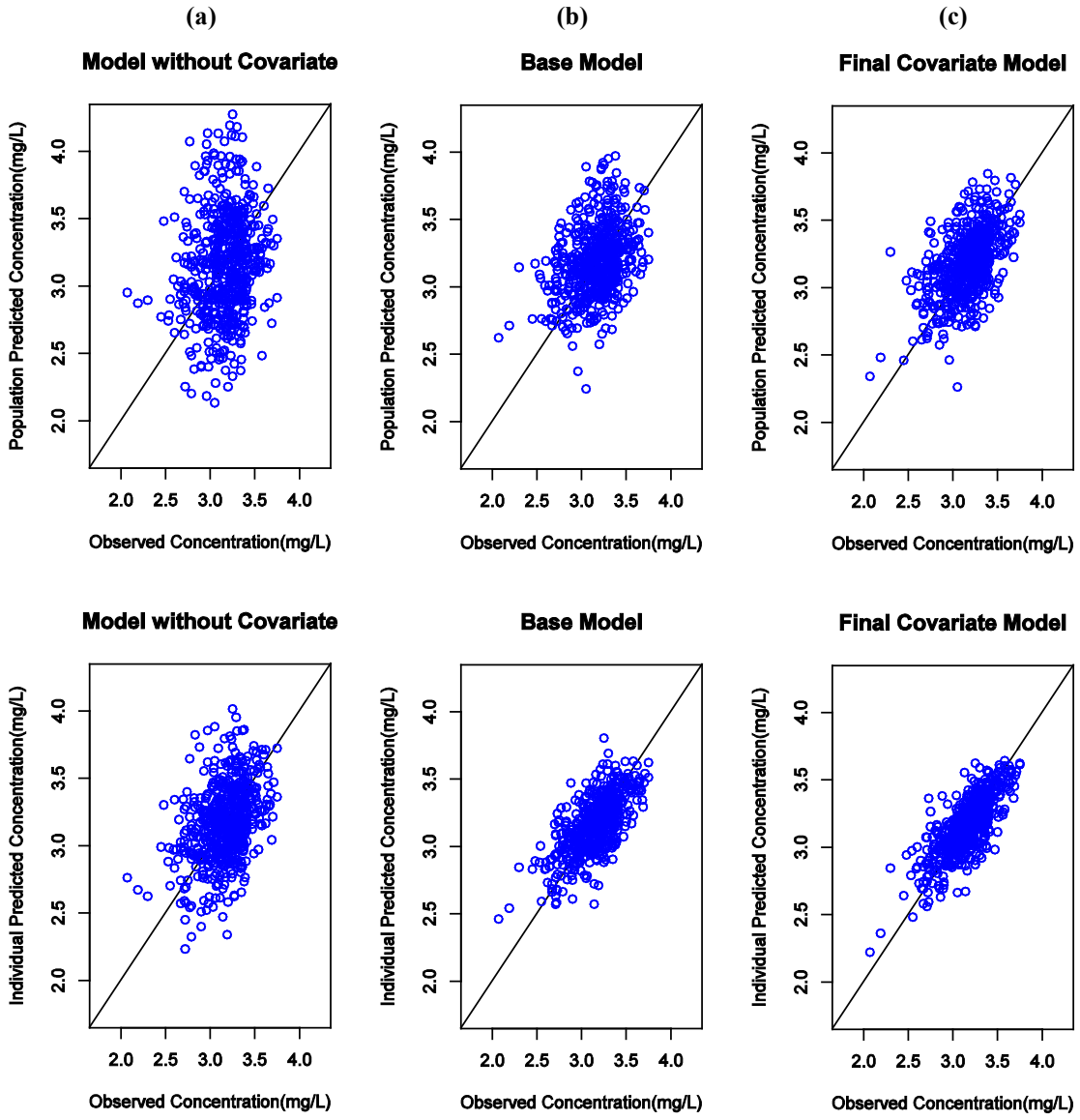


Figure 4-3. Goodness-of-fit plots for the structural model, the base model and the final model

Population (upper panel) and individual (lower panel) predicted versus observed caffeine concentrations obtained from the structural model (a), base model (b), and final model (c).

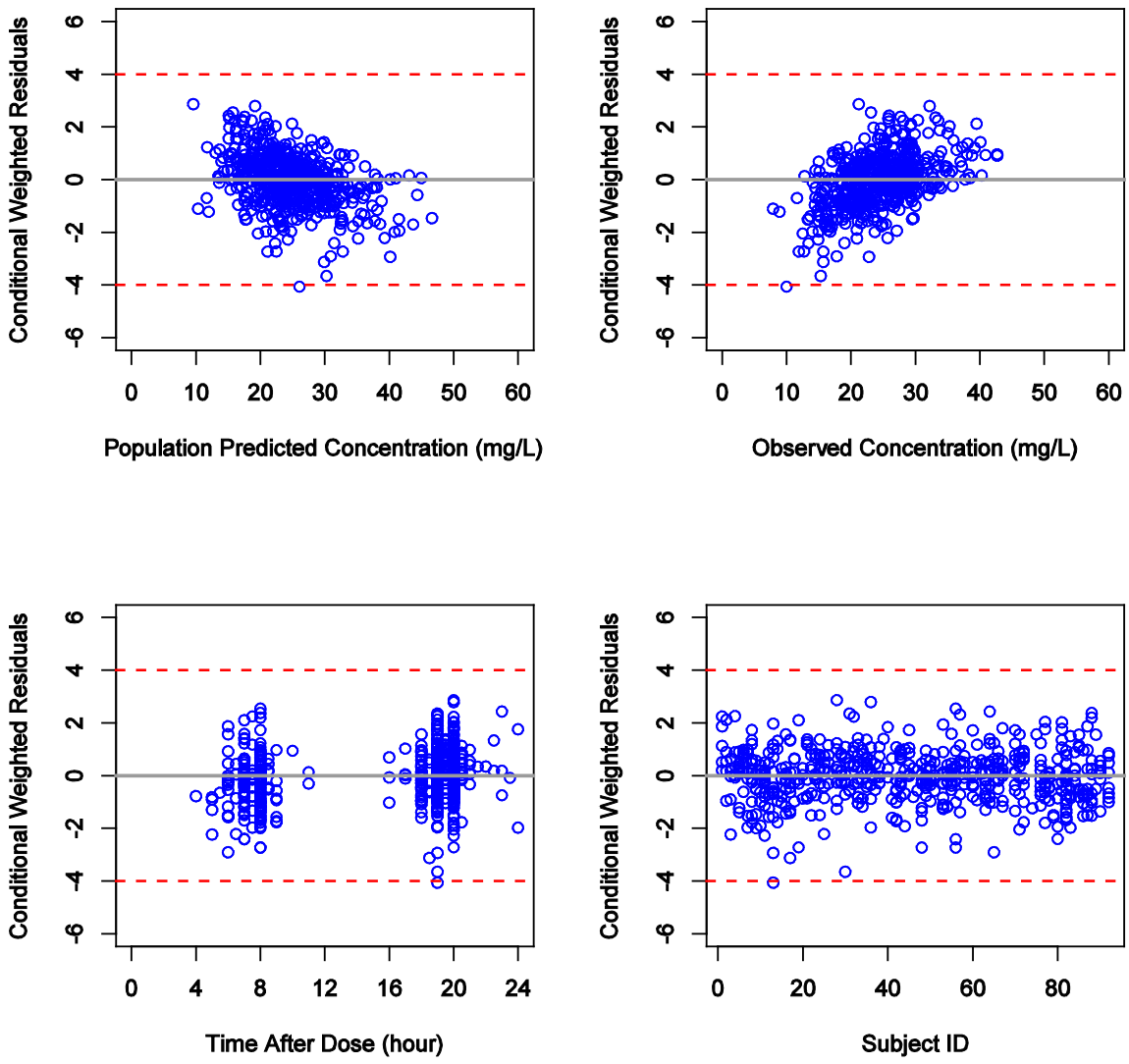


Figure 4-4. Conditional weighted residuals versus population predicted concentrations, observed concentrations, time after dose and subject ID in the final covariate model

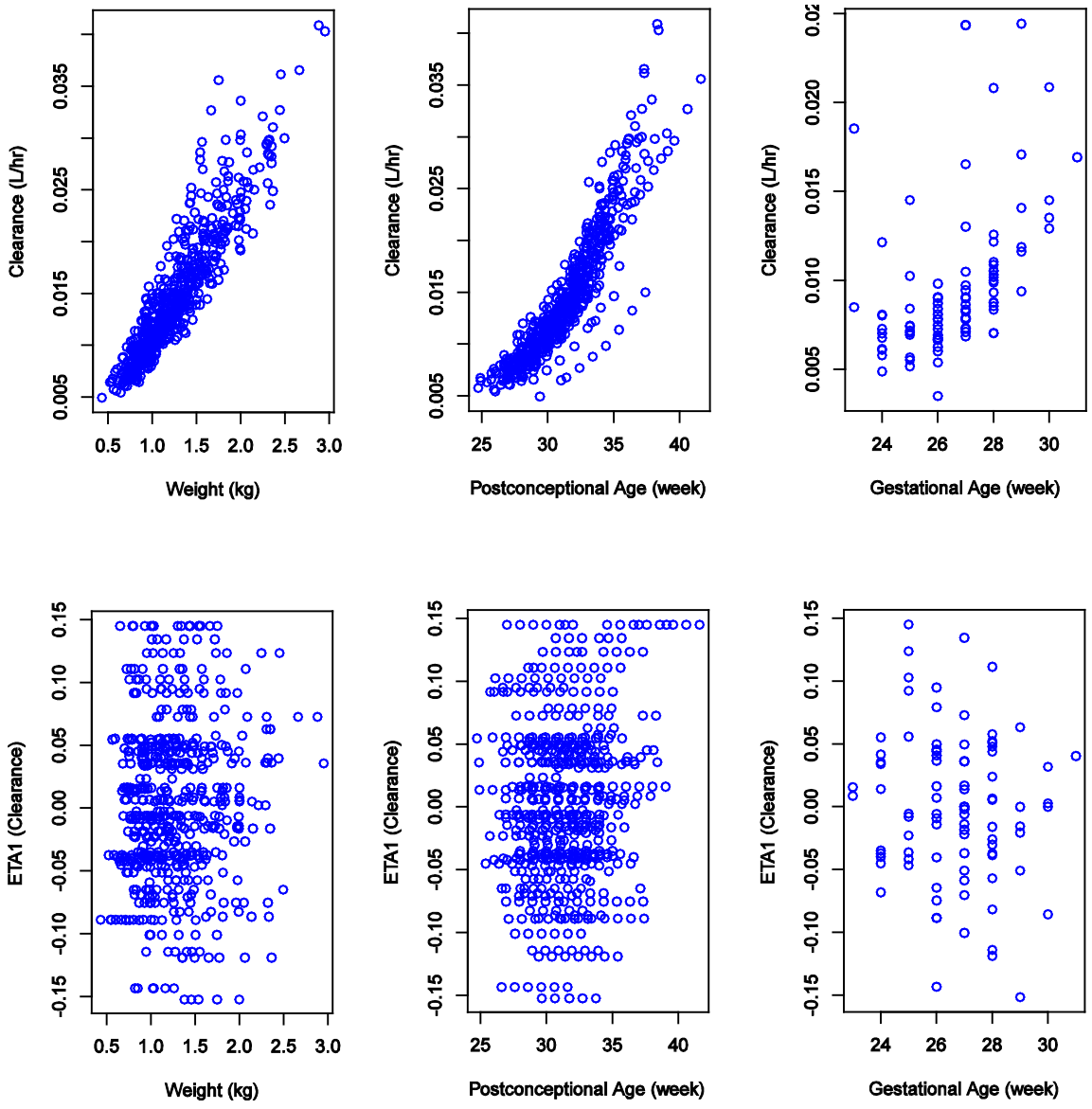


Figure 4-5. POSTHOC individual estimates of clearance and its variance term (ETA1) obtained from the final model versus covariates

Upper panel: scatter plots of POSTHOC individual estimates of clearance versus covariates including body weight, postconceptional age, and gestational age. Lower panel: scatter plots of variance term for clearance (ETA1) versus covariates, including body weight, postconceptional age and gestational age, obtained from the final model.

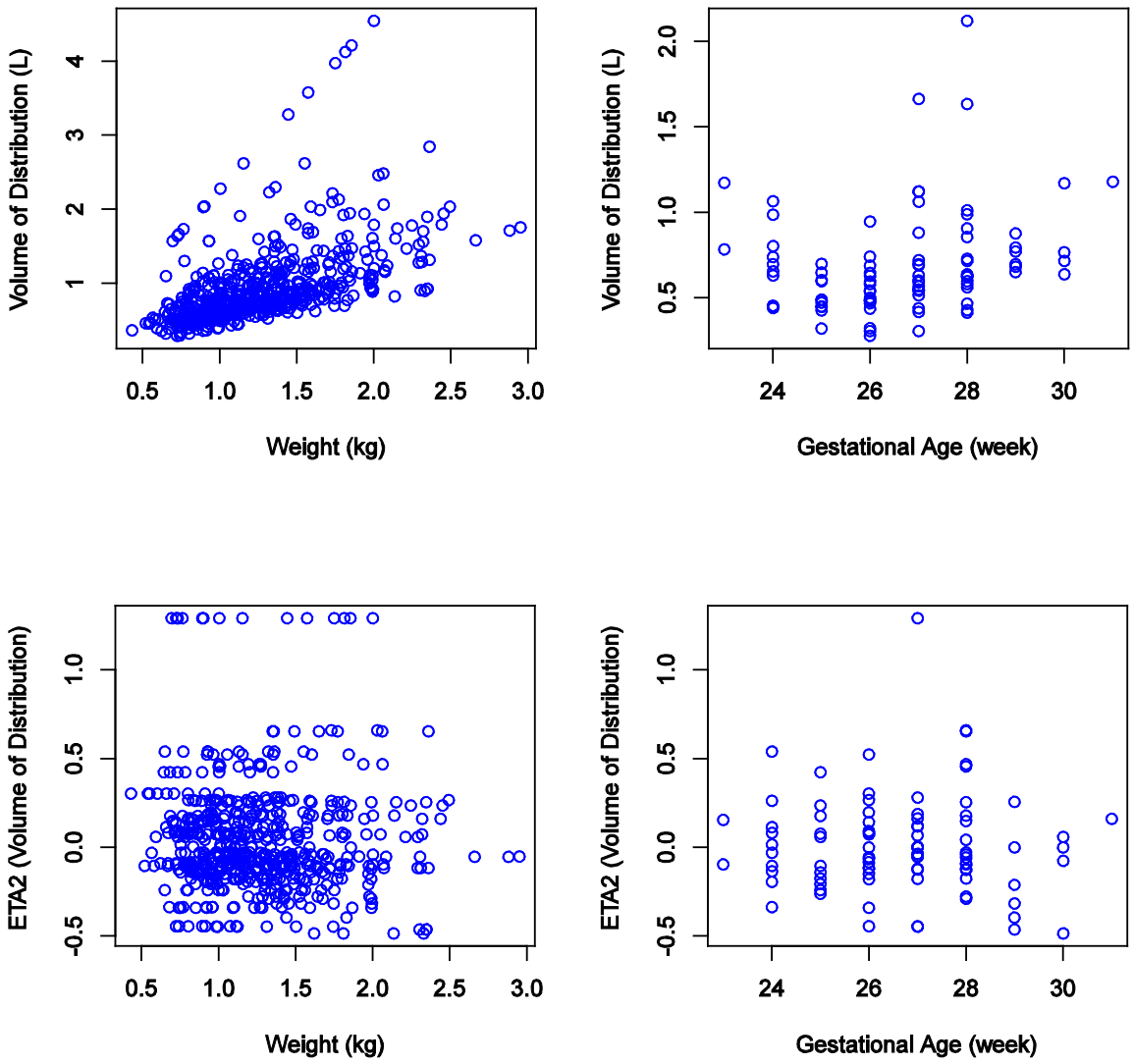


Figure 4-6. POSTHOC individual estimates of volume of distribution and variance term (ETA2) obtained from the final model versus covariates

Upper panel: scatter plots of POSTHOC individual estimates of volume of distribution versus covariates, including body weight and gestational age. Lower panel: scatter plots of variance term for volume of distribution (ETA2) obtained from the final model versus covariates including body weight and gestational age.

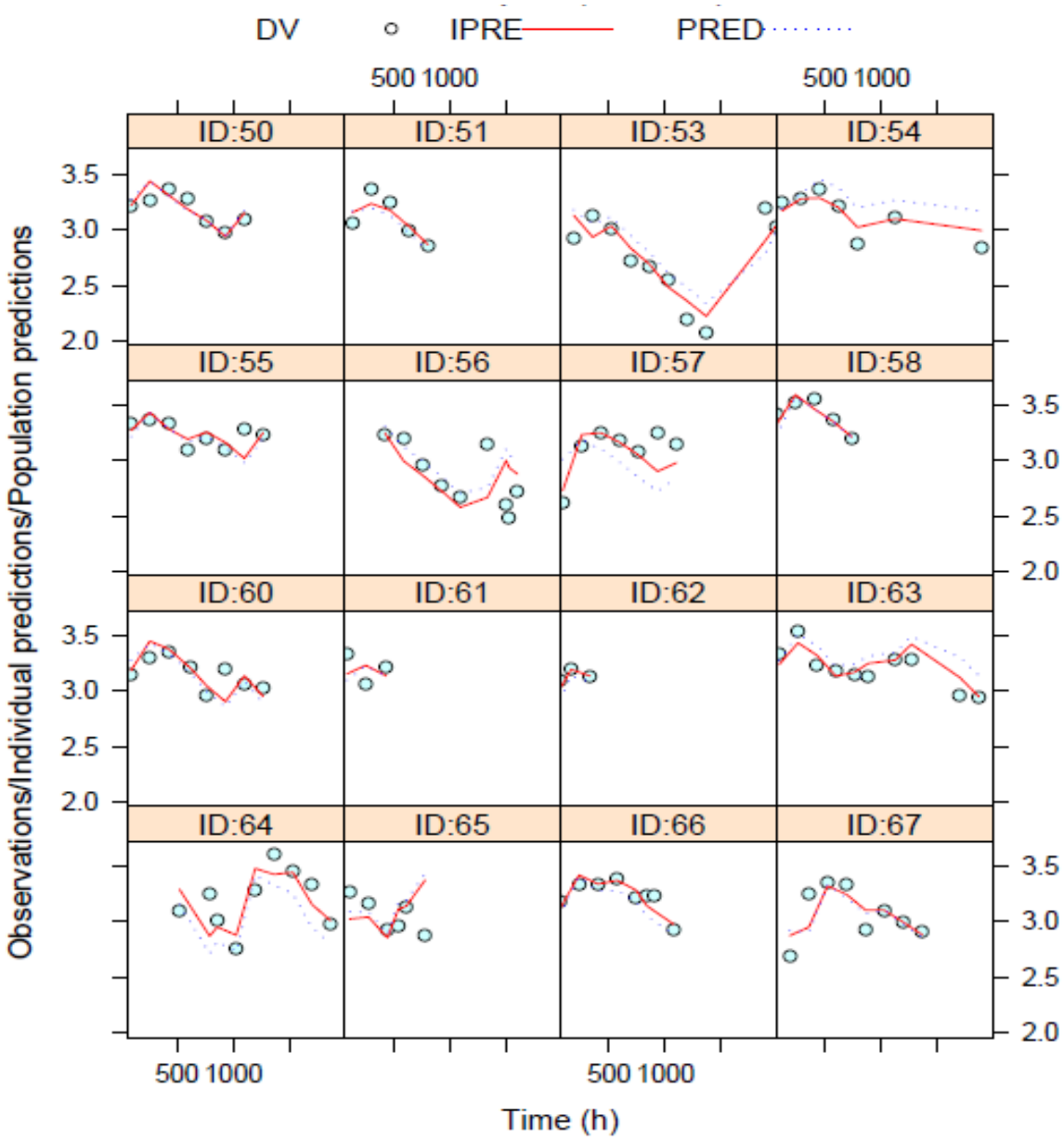


Figure 4-7. Caffeine plasma concentration-time profiles in selected sixteen patients

Patient ID as indicated. Plots of population predicted, individual predicted and observed concentrations versus time. Measured plasma concentrations versus time are indicated by open circles; model-based population and individual predicted concentration-time profiles by dashed and solid lines, respectively.

Model diagnostics

Nonparametric bootstrap

Results from the bootstrap analysis of 500 datasets indicated a stable final model. 100% bootstrap runs attained successful minimization. **Table 4-5** shows that the difference in parameter estimates between the original NONMEM input dataset and the bootstrapped datasets were generally less than 5%, indicating good stability of the final model. Only the estimation of BSV on V ($\sqrt{(\omega_2)}$ x100%) from the bootstrap was 11.5% lower than that from the final model; however, its 95% CI was narrower and completely covered by the 95% CI of the same estimate in the final model. The 95% confidence intervals for the covariate effects did not overlap with 0, indicating the statistical significance of the covariates included in the final model. The 95% confidence intervals were relatively tight, also indicating good precision of all parameter estimates.

Model predictability

Visual predictive check plots were constructed to evaluate the model predictability. **Figure 4-8** shows the observed caffeine concentrations, together with the median and predicted 90% CIs from 500 simulation data sets based on the final model. The visual predictive check confirmed that a majority of the observed values were within the 90% CI of the simulated concentrations by time with similar spread. Therefore, the final full model is considered to be predictive for the model development dataset.

Shrinkage

ETA shrinkage and epsilon shrinkage were calculated for the base model and final model. The shrinkage evaluation results are presented in **Table 4-6**. As mentioned earlier in this section, the dataset used in the PopPK analysis included 88 patients who contributed 1-17 PK observations. While the majority of patients contributed 4-17 PK observations, 13 patients had only 1-3 PK observations. When data are uninformative, the variance for the distribution of individual parameter estimates will shrink towards zero and eta shrinkage will be close to one. Meanwhile, individual predicted concentrations will shrink towards the corresponding observed concentrations and individual weighted residual will thus shrink towards zero, resulting in a high epsilon shrinkage [157]. As expected, some shrinkage was observed for eta and epsilon, which is consistent with the sparse data used in the current analysis. While the value of using individual estimates in the diagnostic plots is considered to decrease in the case of significant shrinkage (eta and epsilon), the population parameter estimation is not affected. Additionally, less reliance on the individual estimate-based diagnostics should be applied in the model building process when a significant shrinkage is observed for eta and epsilon [157]. In this analysis, model selection did not solely rely on the diagnostic plots using individual estimates; instead, multiple model selection criteria were used to determine the best model, as mentioned earlier in the methods section. Meanwhile, other types of diagnostics, including conditional weighted residuals and simulation-based diagnostics, were used to facilitate the selection of an improved model. Thus, the effects of moderate eta and epsilon

Table 4-5. Comparison of parameter estimates after modeling the model building dataset and the 500 bootstrap derived datasets for the final model

Model Parameters	Final Model		Bootstrap		Difference in Estimate[a] (%)
	Parameter Estimate	95% CI	Parameter Estimate	95% CI	
Clearance (CL, L/hr)	0.0164	(0.0157, 0.0171)	0.0164	(0.0157, 0.0171)	0.0%
BSV on CL (%CV)	0.00793 (8.9%)	(0.00364,0.0122)	0.00755	(0.00341, 0.0127)	- 4.8%
Weight on CL (power)	0.750	NA	NA	NA	NA
PCA on CL (power)	1.96	(1.62, 2.30)	1.97	(1.64, 2.33)	0.5%
Low GA on CL (fraction)	1.18	(1.10, 1.26)	1.17	(1.09, 1.25)	-0.8%
Volume of distribution (V, L)	0.940	(0.802, 1.08)	0.926	(0.775, 1.07)	-1.5%
BSV on V (%CV)	0.179 (42.3%)	(0.0394, 0.319)	0.1585	(0.000305, 0.288)	-11.5%
Weight on V (power)	1	NA	NA	NA	NA
Low GA on V (fraction)	1.57	(1.07, 2.07)	1.59	(1.16, 2.29)	1.3%
Absolute bioavailability (F1)	1.00	(0.963, 1.10)	1.03	(0.970, 1.10)	3.0%
BSA on F1 (%CV)	0.0201 (14.2%)	(0.00732,0.0329)	0.0199	(0.00879, 0.0345)	
Residual error	0.0318	(0.0263, 0.0373)	0.0314	(0.0262, 0.0373)	-1.3%

[a] Calculated as (bootstrap value – final model value) / final model value *100.

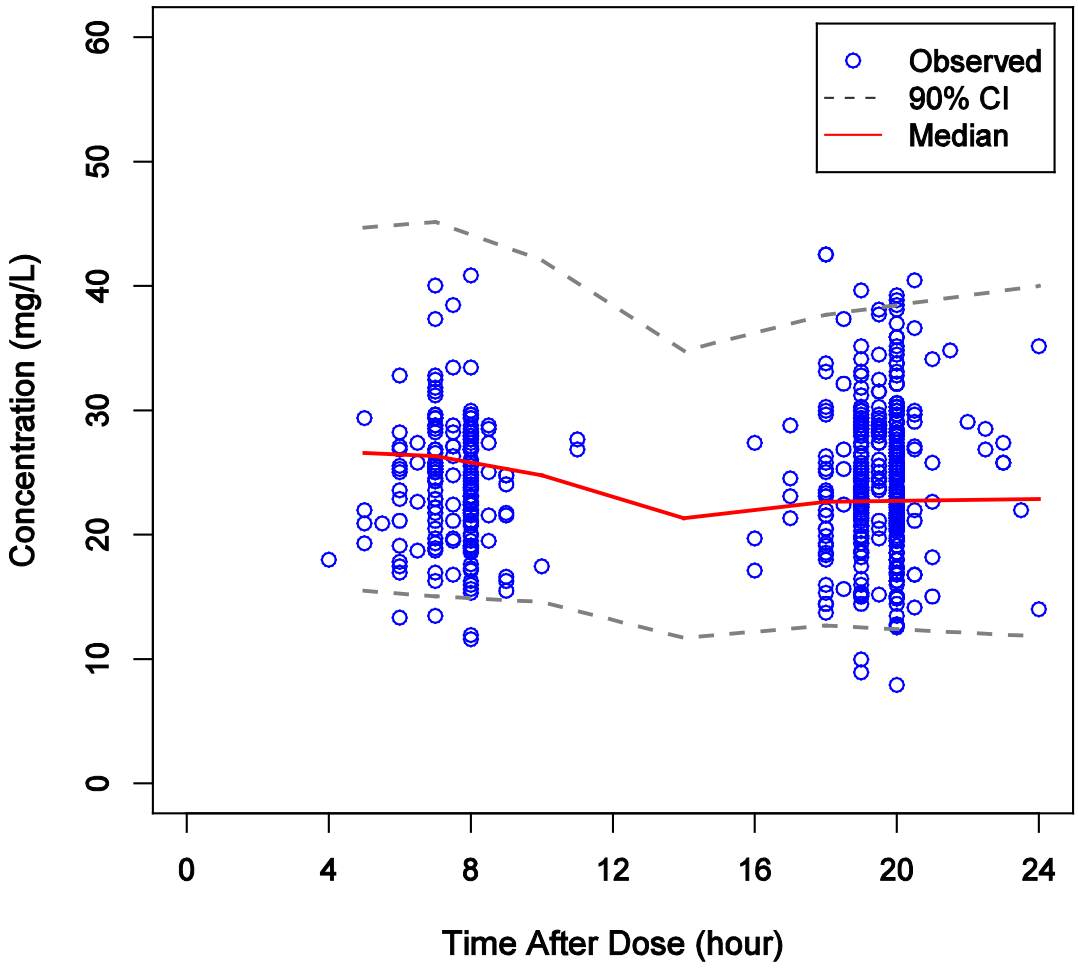


Figure 4-8. Visual predictive check for the final caffeine population pharmacokinetic model.

The red solid line and grey dotted lines indicate the median and 90% confidence interval of predicted concentrations determined from 500 Monte Carlo simulations with the final model. Open circles indicate observed caffeine concentrations in the model building dataset.

Table 4-6. Shrinkage evaluation for the base and final model

Parameter	Base Model Shrinkage	Final Model Shrinkage
ETA1 (CL)	0.209	0.324
ETA2 (V)	0.367	0.338
ETA3 (F1)	0.417	0.356
Epsilon	0.776	0.778

shrinkage on the modeling results are considered to be limited.

Dose-optimization study

A description of the trial simulation cohorts is summarized in **Table 4-7**. The simulated subjects in each cohort had the same demographic characteristics as the infants in the model building dataset of the caffeine PopPK analysis, except for one hypothetical subgroup. Infants with a body weight > 2 kg and a PCA >28 but ≤ 32 weeks were not available in the model building dataset but were considered clinically relevant. Therefore this group was also included in the simulation study. Consistent with the original caffeine dataset, covariate models were defined with a range of 24 to 42 weeks for PCA [mean = 32 (± 3) weeks] and a range of 0.36 to 3 kg for body weight [mean = 1.3 (± 0.45) kg]. A joint distribution between PCA and WT for each studied subpopulation group was derived from the model building dataset for the caffeine population analysis and modeled with a correlation coefficient of 0.36 (range: 0.209 - 0.407). Subjects following the above covariate distribution were generated by using the Trial Simulator. The simulation was stratified so that 20% of the simulated patients had GA < 25 weeks, while 80% simulated patients had GA ≥ 25 weeks.

Model-based simulation was used to evaluate various candidate dosing regimens with regard to a desired therapeutic target concentration. However, there is currently no consensus on the desired caffeine target range in the pediatric community, and thus targets may vary from hospital to hospital and physician to physician. Based on a literature review and the CAFCIT® package insert, the therapeutic targets to attain the satisfactory efficacy and minimize toxicity were defined for the current analysis as: (1) maintaining a trough caffeine concentration at steady state between 8 and 20 mg/L; and (2) maintaining a peak concentration ≤40 mg/L, since caffeine concentrations > 50 mg/L are considered to cause critical toxicity [113, 158]. Maintenance dose recommendations used for posterior check were derived from the frequencies with which the therapeutic targets were achieved in 400 simulated patients for each cohort. **Figure 4-9** illustrates the simulated caffeine peak and trough concentrations sorted by PCA-body weight group and daily dose.

Based on the simulated dose escalation study results using Trial Simulator, the

Table 4-7. Description of the simulated cohorts

PCA (week)	Weight (kg)	LGA [a] (%)	Dose [b] (mg/kg)	Interval (hour)	Duration (day)	Sample Size	PK Target [c]
$23 < PCA \leq 28$	$0.36 < WT \leq 1$ $1 < WT \leq 2$	20	2-12	12 or 24	15	400	$C_{SS,trough}$ 8-20 mg/L $C_{SS,peak} \leq 40$ mg/L
$28 < PCA \leq 32$	$0.36 < WT \leq 1$ $1 < WT \leq 2$ $2 < WT \leq 3$	20	2-12	12 or 24	15	400	$C_{SS,trough}$ 8-20 mg/L $C_{SS,trough} \leq 40$ mg/L
$32 < PCA \leq 36$	$0.36 < WT \leq 1$ $1 < WT \leq 2$ $2 < WT \leq 3$	20	2-14	12 or 24	15	400	$C_{SS,trough}$ 8-20 mg/L $C_{SS,peak} \leq 40$ mg/L
$36 < PCA \leq 42$	$0.36 < WT \leq 1$ $1 < WT \leq 2$ $2 < WT \leq 3$	20	2-14	12 or 24	15	400	$C_{SS,trough}$ 8-20 mg/L $C_{SS,peak} \leq 40$ mg/L

[a] LGA (%): percent of subjects with low gestational age (< 25 weeks) at birth.
[b] Dose (mg/kg): caffeine base.
[c] Steady state is reached after 15 days of caffeine treatment when half-life is less than or equal to 70 hours. $C_{SS,trough}$ and $C_{SS,peak}$ represent the simulated trough and peak concentrations on the day 16.

Figure 4-9. Trough and peak levels of caffeine concentrations at the steady state in dose-finding simulation study

Red solid lines indicate the predetermined therapeutic target range for trough concentration of caffeine: 8-20 mg/L.
Blue dash lines indicate the predetermined therapeutic target for peak concentration of caffeine: 40 mg/L.
Orange dash lines indicate the toxicity level of caffeine: 50 mg/L.
Green box plots indicate maintenance dose level of caffeine base 2.5mg/kg QD, which is equivalent to 5 mg/kg CAFCIT® (caffeine citrate) approved by the FDA for short-term treatment in premature infants with 28-33 weeks GA.

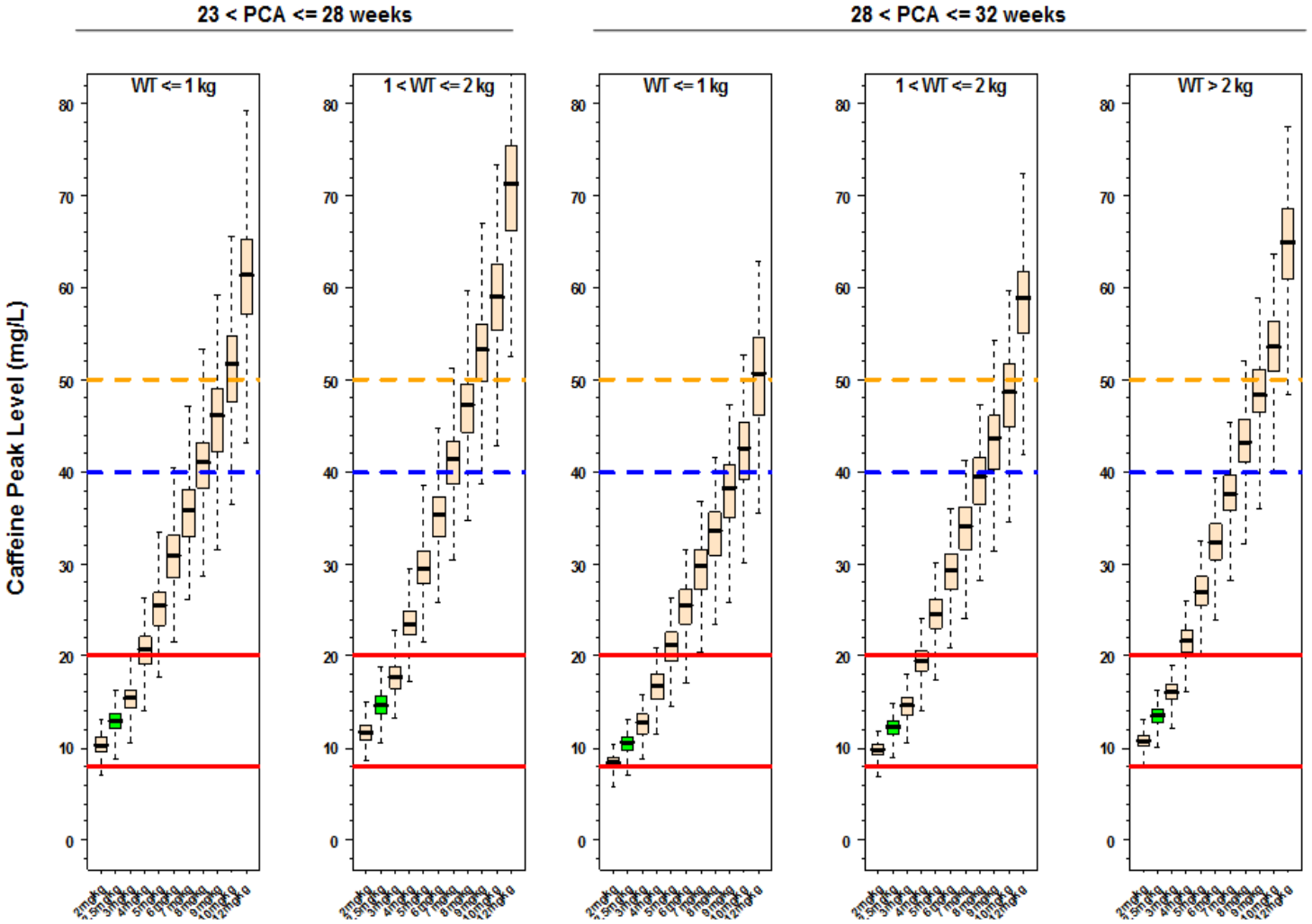


Figure 4-9. Continued

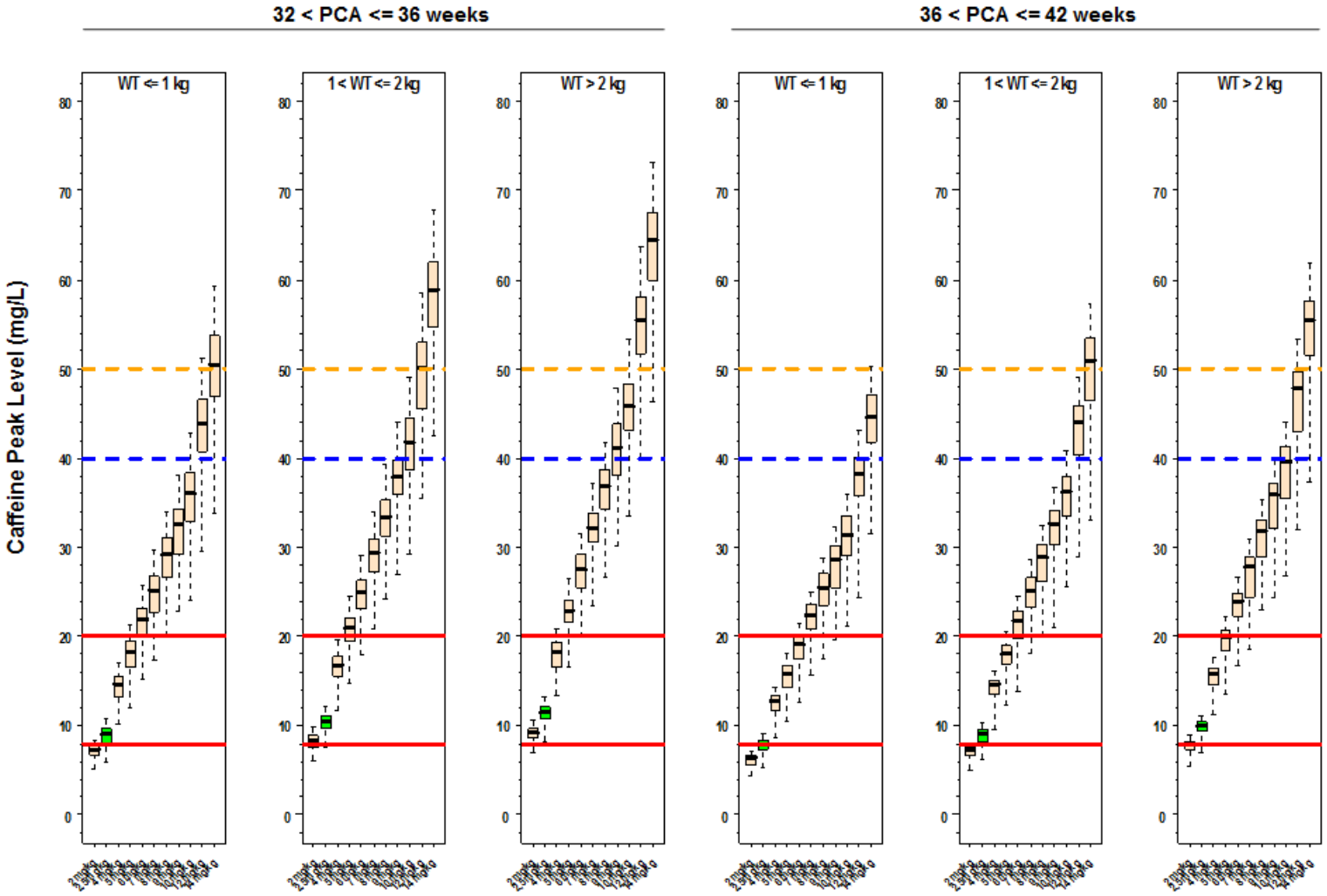


Figure 4-9. Continued

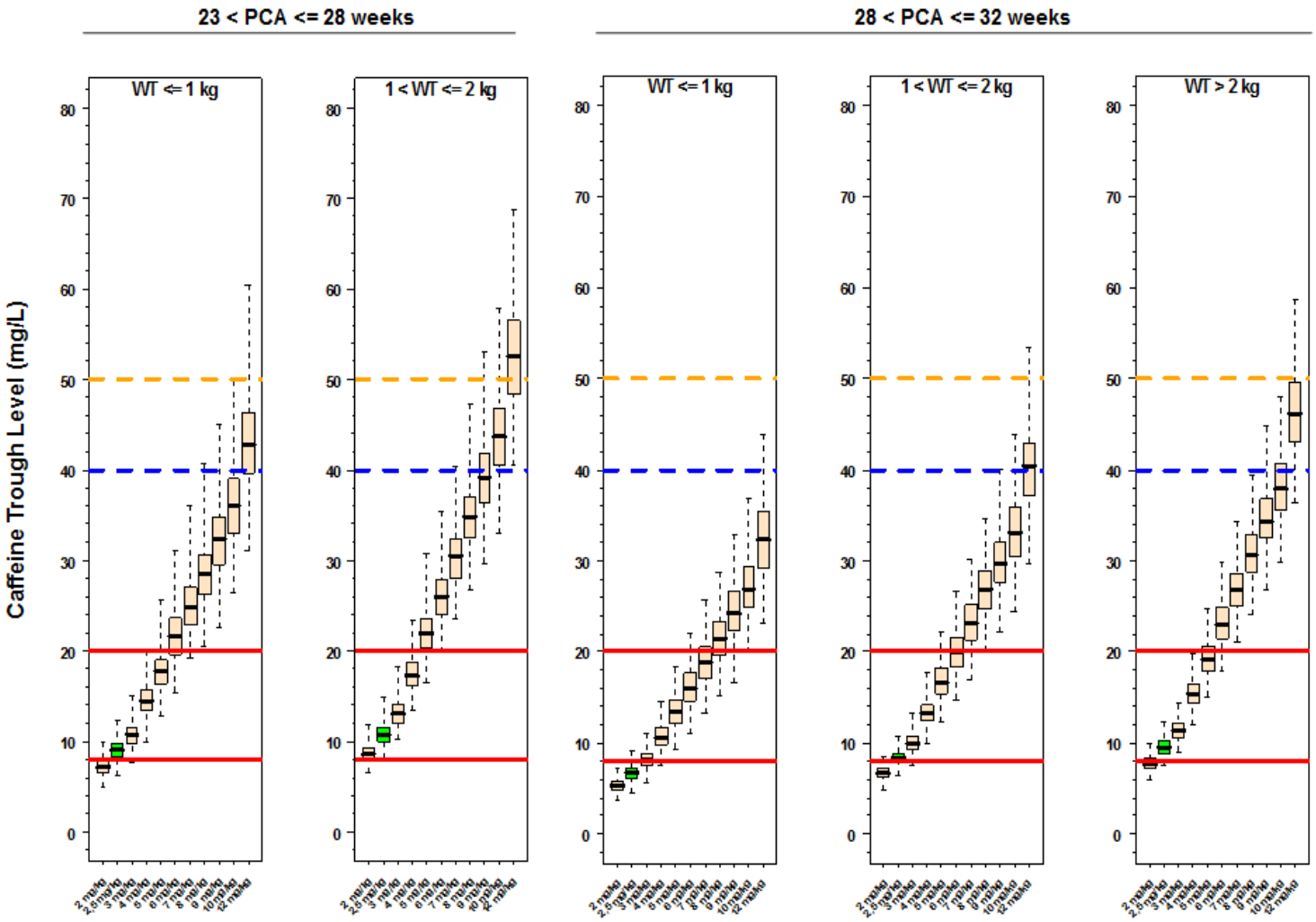


Figure 4-9. Continued

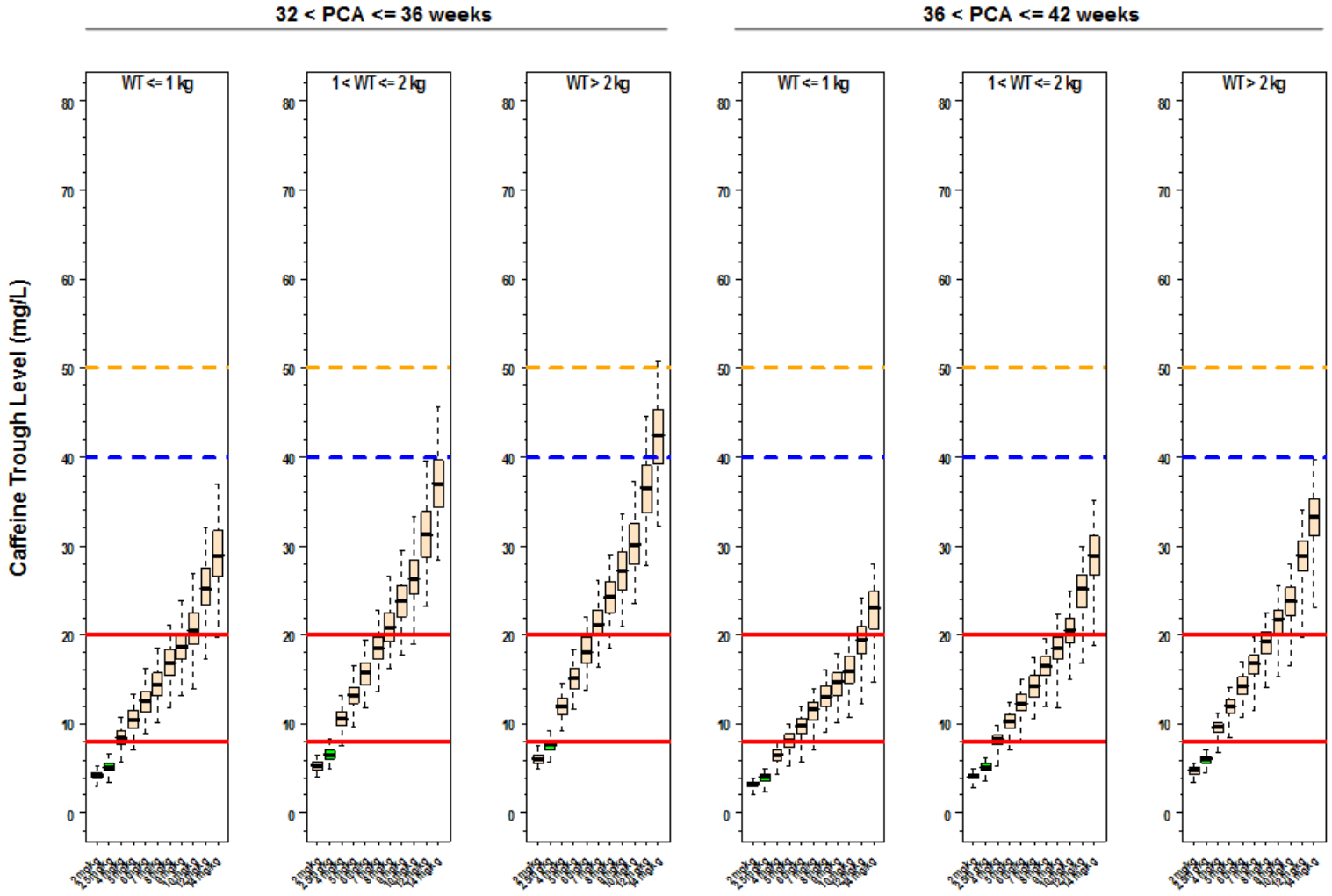


Figure 4-9. Continued

probability of achieving the therapeutic target was calculated for each group and the results are summarized in **Table 4-8**. As illustrated in this table, a higher dose per body weight appeared to be needed to reach the predetermined target concentrations with a decrease of body weight when PCA was > 23 weeks but ≤ 28 weeks. Similar trends were also observed for other three study groups. This was in agreement with the fact that low body weight is usually correlated with low gestational age, and thus may lead to a relatively high clearance per kg body weight for caffeine. Meanwhile, a higher dose per body weight is generally needed with an increase in PCA within the same range of body weight as shown in this figure. Maintenance dose recommendations were derived based on the criterion that the predetermined target range in the simulated premature infants was achieved with a frequency > 99.5%. **Table 4-9** shows that the selected maintenance doses for specific age and body weight groups were consistent with trends as observed in **Table 4-8**. While both 12-hour and 24-hour dosing intervals were evaluated in the dose-finding simulation analysis, dosing every 24 hours can reach the predetermined target and it was selected, since it is more convenient for patient care than a twice daily dosing regimen.

On the basis of the dose-finding simulation results, plasma PK profiles of caffeine in six representative patients after 15 days of dosing were then simulated (n = 500 replicates) under the suggested dosing regimens (**Table 4-9**) using NONMEM. The patients were selected so their covariates covered the range of all studied scenarios in **Table 4-8**, except for the hypothetical group – body weights ≥ 2 kg with a PCA between 28 and 32 weeks, which were not available in the clinical dataset. Simulations were performed using the final PopPK model considering individual parameter estimates, between-subject variability and residual variability. The median predicted concentration-time profiles and 90% CIs from 500 replicates were generated for each patient and compared, with the observed concentrations from the patients included in the Model Building Dataset, who received an empirical and adaptive dosing of caffeine at 5.1-11.7 mg/kg/day, administered every 12 or 24 hours. As presented in **Figure 4-10**, the estimated caffeine concentrations under the new suggested dosage regimen were lower than the observed values. The differences in caffeine concentrations were likely caused by the different therapeutic target concentrations. As discussed earlier, the therapeutic target in the current dose-optimization simulation analysis was a trough caffeine concentration between 8 and 20 mg/L and a steady state peak concentration ≤ 40 mg/L. However, the target trough concentrations for the patients included in the Model Building Dataset were much higher (20-30 mg/L) than the targets used in the current simulation analysis. With the proposed dosing regimens, the predetermined target was well attained, and the simulated median trough plasma concentrations were between 8 and 20 mg/L throughout the treatment period.

The caffeine population pharmacokinetic model

In the present study, we developed a PopPK model that describes the pharmacokinetic characteristics of caffeine in premature infants, covering a GA of 23 to 31 weeks with a PNA of up to 116 days. Body weight, PCA and a low gestational age

Table 4-8. Frequency of Css fall in therapeutic target at the different dose levels of caffeine base

PCA (week)	Weight (kg)	Dose (mg/kg/day)											
		2	2.5	3	4	5	6	7	8	9	10	12	14
23 < PCA <= 28	WT <= 1	24.5	82.5	99.25	99.75	87.25	29.25	1.75	0	0	0	0	NA
	1 < WT <= 2	74.25	100	100	88.75	20.25	0.25	0	0	0	0	0	NA
28 < PCA <= 32	WT <= 1	0	9.25	55	99	100	97	70.5	30.25	9.75	0.25	0	NA
	1 < WT <= 2	6	65.25	98.75	100	93.75	50.75	7.75	0	0	0	0	NA
	WT > 2	32.5	97	100	100	64.5	9	0	0	0	0	0	NA
32 < PCA <= 36	WT <= 1	0	0	NA	65.5	96.5	100	100	94.75	73.5	38.75	4.25	0.25
	1 < WT <= 2	0	2.25	NA	99.5	100	100	78.5	34.75	4.75	1.25	0	0
	WT > 2	0	32	NA	100	100	82	24.25	2.25	0	0	0	0
36 < PCA <= 42	WT <= 1	0	0	NA	0	54	92.25	99	100	100	100	57.75	6.25
	1 < WT <= 2	0	0	NA	63.75	98.25	99.75	100	100	78.75	38.25	3.5	0.25
	WT > 2	0	0	NA	94	100	100	100	67	22	4.25	0.25	0

PCA = postconceptional age; WT = weight; NA = not applicable.

Table 4-9. Suggested intravenous maintenance dose (mg/kg/day QD) for caffeine base in premature infants stratified by PCA and body weight

PCA (week)	Body Weight (Kg)		
	< = 1	(1,2]	> 2
< = 28	4	3	
(28, 32]	5	4	3
(32,36]	6	5	4
> 36	8	6	5

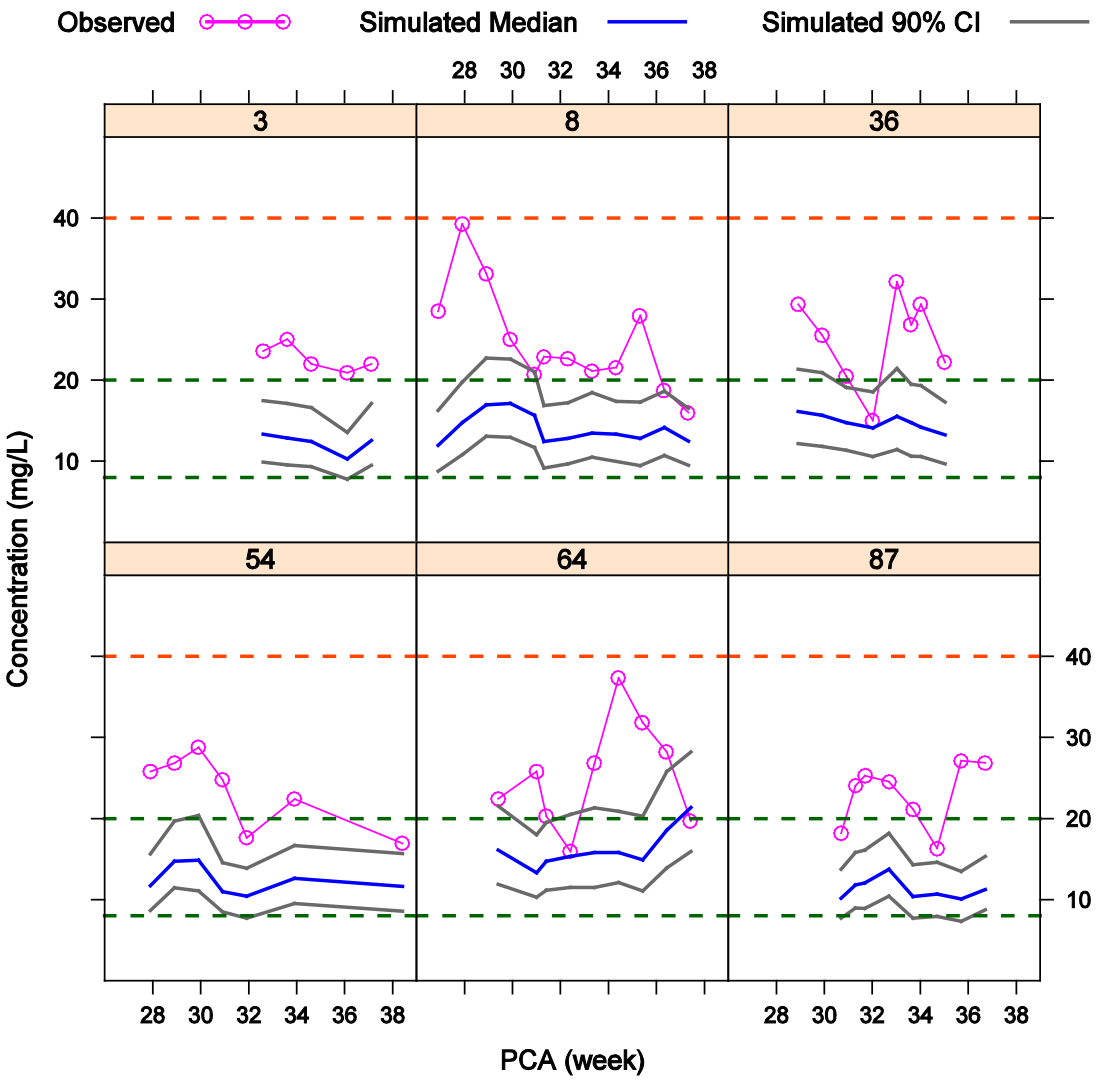


Figure 4-10. Simulated caffeine concentration-time profiles in 6 representative patients using the dosing regimens suggested in Table 4-9.

Observed caffeine plasma concentrations after adaptive dosing are indicated by open circles; blue and gray solid lines represent the median and 90% confidence intervals generated from 500 Monte Carlo simulations.

< 25 weeks were found to be important predictors explaining the between-subject variability of caffeine pharmacokinetics in premature infants receiving caffeine treatment. Overall, the BSV of CL decreased from 28.0% to 8.9% with the addition of body weight, PCA and low GA factor on CL. Particularly, inclusion of body weight, PCA and low GA factor reduced the between-subject variability in CL from 28.0% to 14.9%, from 14.9% to 10.3%, and from 10.3% to 8.9%, respectively. The median (range) empirical Bayesian estimates of parameters for all individuals were 11.6 (7.1-20.3) mL/hr/kg for clearance and 0.67 (0.39-2.28) L/kg for volume of distribution. These results are in agreement with those reported in the literature involving premature infants receiving caffeine treatment (**Table 4-10**).

An allometric model with fixed exponents of 0.75 and 1 was used to model the effect of body weight on clearance and volume of distribution. The allometric principle has a concrete ecological rationale based on fractal geometry theory and is robust when used as size adjustment in pharmacokinetics [39, 159]. This rationale allowed us to explore the unique effect of other covariates, despite their collinearity with body weight [73]. Body weight explained 46.8% of the reduction in BSV for CL. PCA was subsequently found to be another significant covariate of caffeine's CL, and clearance increased nonlinearly with PCA. The addition of PCA on CL explains 30.7% of the reduction in BSV for CL. Both PCA and PNA were determined to be significant predictors for CL when tested stepwise in the base model during the model building procedure. The addition of PCA on CL demonstrated a better improvement in the model fit and explained more of the BSV of CL than PNA did. After the addition of PCA on CL, PNA did not improve the model fit further and resulted in over parameterization. This result is also in accordance with our hypothesis prior to the covariate screening. Compared to PNA, PCA is considered more physiologically appropriate [160], especially when maturational processes of CL are initiated before birth and/or the study population is heterogeneous with respect to GA and PNA at the onset of the pharmacotherapeutic intervention. In this case, PCA was expected to contribute more relevant information with regard to maturation of drug disposition processes rather than PNA or GA alone.

A sigmoid maturation model was also investigated to describe the age effect on clearance maturation since the maturation process of CYP1A2 could be modeled by this Hill type function [161], as seen in the following equation:

$$\text{Maturation Function} = \frac{(A)^n}{A_{50}{}^n + (A)^n}$$

Where A represents the age term, such as PCA or PNA; A_{50} is the age term at which clearance is 50% that of the mature value and n is the Hill coefficient. Both PCA and PNA were tested but did not result in a successful model fit. There are several reasons that might account for this failure. First, this model might better describe those maturational progresses triggered by parturition [39]. However, the maturation of processes relevant for caffeine's clearance may start already before birth. Secondly, there is no clear information available for the ontogeny of clearance mechanism in the studied population. The elimination of caffeine in neonates is not dominated by one pathway,

Table 4-10. Comparison of pharmacokinetic parameters of caffeine in premature infants in our study and as reported in the literature

Reference	N	CL (mL/h/kg)	V (L/kg)	$T_{1/2}$ (h)	Birth Weight (kg)	GA (week)	PNA (day)	PCA (week)	Weight (kg)
Current study	88	11.6 (7.1-20.3)	0.67 (0.39-2.28)	40.0 (16.8-162.1)	0.84 (0.38-1.7)	26 (23-31)	39 (1-116)	32 (24-42)	1.3 (0.36-3.0)
Aranda et al. [153]	12	8.9 (2.5-16.8)	0.92 (0.48-1.28)	102.9 (40.8-231)	1.11 (0.69-1.87)	28.5 (25-34)	11.5 (3-32)		
Gorodischer et al. [162]	13	8.5 (5.8-12.2)	0.78 (0.47-1.01)	65 (48.2-87.5)	1.4 (0.92-2.06)	30.6 (25-34)	1-42		
Le Guennec et al. [160]	23				1.42 (0.64-2.35)	30 (25-36)			
Thomson et al. [163]	60	7.9	0.82				23 (1-100)	31 (25-41)	1.3 (0.6-2.9)
Falcão et al. [151]	75	7.6	0.91		0.6-2.0	23-35	1-78	26-38	0.6-2.7
Lee T.C. et al. [164]	89	4.4-5.6	0.86-1.11		1.17 (0.57-2.31)	28.2 (24-31)			
Micallef et al. [165]	35				1.34 (0.66-2.17)	29.1 (23-32)			
Saleh Al-Alaiyan et al. [166]	80	7.62 (2.8-30.2)	NA	NA	1.3 (0.65-2.26)	30 (24-34)	28 (5-60)	34 (29-40)	1.63 (0.98-2.67)
Lee H.S. et al. [167]	18	6.28 * (17.5%)	0.96 ** (20.3%)		1.12 (0.68-1.7)	28.9 (24-33)	NA	NA	NA
Charles et al. [152]	110	6.96 (1.61-22.6)	0.85 (0.37-1.76)	101 (24.5-371)	1 (0.568-1.57)	27.6 (24-29)	12 (1-45)	29 (24-34)	0.99 (0.66-1.86)

Table 4-10. Continued

* unit (L/h) ** unit (L)

instead, complementary pathways keep changing as the maturation of kidney and liver proceed, resulting in a complex effect on the overall CL. Thirdly, variation of other factors relevant for the disposition of caffeine, such as low albumin levels (binding protein), body composition or bilirubinemia may confound estimations of hepatic clearance [168, 169]. Another possible reason might be the limited data in this study: They are not rich enough to support estimating parameters in this sigmoid maturation model and age (PCA, PNA) span in the analysis dataset is too narrow to characterize this maturation function appropriately.

Study results compared to the literature

Compared to the published literature, our study presents substantial new findings in those extremely premature infants with GA < 25 weeks. A patient with GA < 25 weeks at birth has 18% higher allometrically weight adjusted CL and 57% higher weight adjusted V compared to one with GA > 25 weeks. These findings suggest that a relatively higher loading dose and higher maintenance dose based on body weight would benefit in treating AOP in premature infants with a GA < 25 weeks at birth. The relatively larger estimate for V is also consistent with estimates previously reported, where body weight normalized V for caffeine in premature neonates is larger than that in term neonates and adults [101, 170]. Similar dose guidance was suggested by a controlled trial of caffeine citrate [101], when Erenberg et al. proposed a larger loading dose for less mature infants [101]. It is probably because there is proportionally more extracellular fluid in less mature infants, which leads to an age-dependent larger water/body weight ratio [171]. Caffeine, as a highly hydrophilic drug, would be expected to demonstrate an increased apparent volume of distribution under these conditions [26, 48]. It should be noted that a relatively large BSV (42.3%) on V was observed in the current study. This is likely due to the fact that most concentration data were collected at trough level during the TDM. Little information is available for the estimation of the absorption phase and peak concentrations, thus affecting the estimation for V. Another reason might be that water loss could be easily induced in premature neonates by environmental factors or clinical treatment, such as phototherapy, using a radiant warmer, or diuresis after caffeine treatment. As a result, a large between-subject variability and within-subject variability for V is often observed in this patient population.

The inclusion of a low GA factor in the CL model indicates that size and PCA alone cannot explain well the maturation progress of caffeine CL in extremely premature infants. PNA might play an important role in this process as well, where the postnatal development trajectory does not follow the intrauterine curve during the early postnatal life. A literature review also showed that most previous caffeine studies had detected a correlation between caffeine CL and PNA [152, 163, 166], PCA [160] or the combination of GA and PNA/PCA [151, 164]. In the studies by Falcão [151] and Lee [164], a corrected factor of GA ≤ 28 weeks was incorporated in the CL model, which found that PNA and weight had influence on CL. It has been reported by Pons et al [172] that the development of caffeine clearance after birth reached the plateau after 4-6 months, while a linear relationship was observed with gestational age and an exponential relationship

with postnatal age based on a study with a PNA range of 15~588 days. These findings are in agreement with our conclusion that together with size as a covariate, neither PNA nor PCA can solely characterize the maturation process of caffeine disposition in premature infants. That process might be described better by the combination of two types of age covariates, as implemented in our population model. In this way, the variable degree of maturation at birth and postnatal development are both accounted for in the CL model. This approach is also supported by a fluconazole population study in neonates and infants, where both GA and PNA were included in the CL model for fluconazole [173].

The ontogeny of caffeine elimination

Understanding the impact of the ontogeny of various elimination pathways on caffeine CL in premature infants remains incomplete. The total clearance of caffeine in premature neonates cannot be compared to term neonates and adults due to the immature liver and kidney functions at birth. In adults, caffeine is predominantly metabolized through cytochrome P450 (CYP) enzymes in the liver, and only approximately 2% of caffeine is excreted unchanged in urine [154, 174]. Of the human CYP enzymes investigated, demethylation through CYP1A2 plays the most important role [175-178]. Besides, CYP2E1, CYP3A4, N-acetyltransferase (NAT) and xanthine oxidase (XO) are all associated with caffeine's biotransformation as well [166, 177, 179, 180].

In contrast to the adult situation, the renal pathway is thought to be compensatory in newborns and remains dominant for at least 3 months after birth. As indicated by a lack of expression of mRNA in human fetal liver, there is no significant enzymatic activity of CYP1A2 at birth [181]. Development is believed to be triggered by parturition [166, 181]. Previous studies support these mechanistic considerations by reporting that renal clearance plays an important role on caffeine's elimination in neonates. Transplacentally acquired caffeine was found to be almost completely recovered in urine during the first 3 days after birth [182]. Additionally, 85% of the ingested dose was excreted unchanged in the urine during the first month of life [174]. However, renal clearance is less efficient than the CYP1A2-mediated metabolism pathway. Thus, the magnitude of the reduction of overall caffeine clearance in neonates could be as large as 10-fold [170]. Adult clearance levels are reached in 4 to 6 months after birth [48, 130, 160, 183], which reflects largely the metabolic activity of CYP1A2.

Nephrogenesis starts as early as 9 gestational weeks. Neonatal glomerular function shows a progression positively correlated to GA and PNA [31, 184, 185]. There are data to show that renal creatinine clearance (CrCL) was significantly lower in the low GA population (GA < 30 week), and steadily increased after birth [184, 186]. These findings about the development of renal elimination in neonates support the selection of GA and PCA as predictors for caffeine CL. Additionally, a considerably decreased glomerular filtration rate (GFR) was detected in neonates with perinatal asphyxia [187]. Although creatinine values were significantly higher in preterm babies than in term babies in the first week, they reached almost similar levels by the third week of life [188]. Such data indicate that a relatively faster development of renal function might be

associated with less mature neonates during the first 3 weeks of life, and thus may explain the effect of a low GA on caffeine CL. However, further studies are needed to clearly define the underlying mechanism of the development of caffeine elimination pathways.

Dosing recommendations

An essential goal of the current dose-optimization simulation study was to find appropriate dosing regimens that can reach a trough level of 8-20 mg/L and a peak level ≤ 40 mg/L for caffeine at steady state for the treatment of AOP in premature infants. Using trial simulation, four clinically relevant age categories were investigated using doses from 2 mg/kg to 14 mg/kg with a 12-hour or 24-hour dosing interval. Although patients received maintenance doses ranging from 3.1 to 28.6 mg/kg/day with either 12 or 24 hours of dosing interval in the Model Building Dataset, a 24-hour-dosing interval was shown to be successful by the Monte Carlo simulations in the present study with respect to the proposed target concentrations in all simulated groups. This finding is also supported by the remarkably prolonged half-life of caffeine as reported previously in this population [153, 160]. However, a trend was noted that peak-trough fluctuation at steady state appeared to increase with the increase of patients' body size and age.

Moreover, a large variability on the clinical responses of caffeine was observed in the original Model Building Dataset, where an empirical and adaptive dosing was used with the target trough concentrations between 20 and 30 mg/L. For example, 13 out of 88 patients still had uncontrolled apnea and bradycardia episodes (1-26 episodes/patient) following at least 10 days of caffeine treatment. Meanwhile, 50 out of 88 patients had tachycardia (heart rate > 180 bpm) during the caffeine treatment period, where caffeine trough concentrations ranged from 7.9 to 42.7 mg/L. This variability is likely due in part to limitations of the empirical dosing strategy, where the therapeutic target is adjusted individually through TDM to achieve an effective clinical effect. However, the dose-optimization simulations based on the developed PopPK model may provide more benefit while allowing clinicians to compare various dosing regimens and bridge the plasma caffeine levels with response at different PCAs and different body weights. Therefore, dose selecting for alternative targets would be more confirmatory than explorative.

A large variability was estimated for volume of distribution of caffeine from the final PopPK model. Since most concentration measurements were collected at trough level of each dosing interval and also at least 40 hours after the loading dose, information regarding the estimation of volume of distribution is thus limited. As a result, the optimization of the loading dose was not evaluated in this simulation analysis.

Conclusion

In conclusion, a PopPK model of caffeine was developed for premature infants: body weight, PCA and information of gestational age were identified as important

predictors of variability of caffeine pharmacokinetics. Furthermore, we investigated the application of this PK knowledge to further facilitate the development of optimal dosing regimens through trial simulation. Various dosing regimens with a dosing interval of 12 or 24 hours were evaluated to reach the predetermined therapeutic target with a trough level of 8-20 mg/L and a peak level ≤ 40 mg/L at steady state. A dosing interval of 24 hours was shown to be successful with respect to the proposed target concentrations in all simulated groups. While the modeling and simulation approach is not intended to replace the TDM, it may provide valuable reference and information when dose adjustment becomes clinically urgent, especially for those patients with various maturational levels at the initiation of therapy. The successful use of modeling and simulation approaches in neonatal and infancy studies may reduce the number of invasive blood drawings required by therapeutic drug monitoring. Moreover, the posterior predictive check showed successful dose adjustment within an individual over time. A rational dosing regimen could be determined more rapidly based on the PopPK characteristics of caffeine rather than by empirical dosing.

CHAPTER 5. SUMMARY

Currently, "off-label" use of drugs in pediatric populations remains an acknowledged problem across almost all categories of therapeutics throughout the world [2-9]. It has been reported that up to 62% of pediatric outpatient visits involve off-label or unlicensed medications [3]. Another report indicated that 70% of the medications in pediatric intensive care and 90% of the medications in neonatal intensive care were given in an off-label manner [10]. Off-label prescribing is more common in younger pediatric populations compared to the older children, especially in premature neonates–the primary group receiving intensive care [3, 4], and dosing regimens for most drugs used in neonates are usually empirical. Limited clinical data is considered the major reason for the "off-label" use of drugs in pediatric populations, due to the complexity of pediatric studies as well as scientific, logistical and ethical concerns. Major pharmacokinetic challenges to assuring proper pharmacotherapy in premature neonates and infants are: limited volume and frequency of blood sample collections, rapid growth and continuous developmental changes and empirical dosing due to the lack of pharmacokinetic information. My dissertation research focused on several approaches to overcome these unique challenges in premature neonates and infants. One of them was to develop an accurate and sensitive LC-MS/MS assay, which can simultaneously quantitate multiple drugs frequently used in pediatric pharmacotherapy using a small volume of plasma. Additionally, PopPK modeling and simulation for sample size estimation and appropriate study design and dosing regimen were investigated to improve drug development and pharmacotherapy in pediatric populations.

LC-MS/MS is a standard bioanalytical methodology for drug research due to its robustness and high sensitivity, which allows for reliable quantification even within the confines of small sample volumes in pediatric studies [104]. The first objective of my research was to develop an LC-MS/MS assay for the simultaneous determination of acetaminophen, caffeine, phenytoin, ranitidine and theophylline using small volume human plasma specimens for pharmacokinetic evaluation. These five drugs were selected as they are all currently widely used in the pharmacotherapy of premature and term neonates [98-103], with only limited pharmacokinetic information available. Due to the limitations in sample volume, developing an assay that can simultaneously determine multiple drugs allows for gaining maximal information from pharmacokinetic studies while minimizing the burden on pediatric patients. An accurate, sensitive, and reliable LC-MS/MS method was developed and validated using small volume of 50 µL human plasma to quantitate the selected five drugs simultaneously with mean accuracy ranging from 87.5% to 115.0%, and intra-day and inter-day precision ranging from 2.8% to 11.8% and from 4.5% to 13.5%, respectively. This assay quantifies a range of 12.2 to 25,000 ng/mL for acetaminophen, phenytoin and ranitidine, a range of 24.4 to 25,000 ng/mL for theophylline, and a range of 48.8 to 25,000 ng/mL for caffeine. These ranges cover each drug's therapeutically used concentrations in the neonatal age group. A sample dilution procedure was also evaluated, and the results indicated that the assay's intra-batch accuracy and precision were not affected by the 1-to-2 dilution. The effects of hemolysis, lipemia and hyperbilirubinemia were subsequently evaluated, and no

interference in the analysis was noted when these factors existed separately or combined. Additionally, no significant matrix effect was observed for the developed bioanalytical assay.

Based on the fact that clinical research in the neonatal population can only be performed within the context of therapeutically necessary interventions, an opportunistic sampling approach was proposed for PK studies—that is utilizing the leftover of blood samples taken for routine clinical care in PK studies. This design is thought to be ethical since no extra invasive blood draw will be imposed on the patients. Successful application of population-based M&S can be used to determine the appropriate number of subjects needed in pediatric studies, thus resulting in fewer patients exposed to the investigational drugs with an adequately powered study. The effect of sample size on the robustness of population pharmacokinetic parameter estimates and covariate detection in the observational study design in premature neonates was evaluated using a full model-based simulation approach with theophylline as the model drug. Simulated datasets for each sample size (9-200 subjects per study) with a mixed and unbalanced sampling design were first generated with the incorporation of changes in birth weight, body weight and PNA. The median PopPK parameters for theophylline estimated from the simulated datasets were generally in close agreement with those of the originating model across all tested sample sizes; while the accuracy, precision and power of parameter estimation benefit from increases in the number of study subjects. The power of the study was deeply influenced by the sample size, parameter of interest and the selected precision level. Furthermore, the power to detect the potential covariate effect was investigated at three significance levels. It was found if the desired bias in parameter estimations in terms of %MPE was $\leq 15\%$, and $\leq 25\%$ and $\leq 50\%$ were accepted as being precise for fixed effect and variance parameter, respectively, a sample size of 40 subjects would be sufficient. At a sample size of 40 subjects, the power to detect the covariate effect was greater than 80% at a significance level of $P = 0.01$. Overall, this proposed approach can also be applied to evaluate the impact of other design factors which may influence the required number of subjects in PK studies in premature infants, such as different sampling design (sparse or dense sampling), allocation of sampling times, estimation methods, and magnitude of variability. It may also prove valuable in studying other drugs of interest if appropriate prior knowledge is available.

There is currently limited PK data on caffeine in premature neonates. We developed a PopPK model of caffeine in premature neonates and identified potential sources of variability of PK behavior for caffeine. We subsequently investigated the application of this PK knowledge to further facilitate the development of optimal dosing regimens through simulation, particularly to correlate steady state concentrations with different dosing regimens in various age/body size groups. In the present study, a one-compartment model was chosen to describe the pharmacokinetic characteristics of caffeine in premature infants, covering a gestational range of 23 to 31 weeks with an age of up to 116 days. Body weight, PCA and a low gestational age < 25 weeks were found to be important predictors explaining the between-subject variability of caffeine pharmacokinetics in premature infants receiving caffeine treatment. Particularly, the BSV of CL decreased by 68% (from 28% to 8.9%) with the addition of body weight, PCA and

a low GA factor on CL. The typical patient in the studied premature neonate population (WT of 1.5 kg, PCA of 32 weeks and with a GA > 25 weeks) was estimated to have a CL of 0.0164 L/hr and a V of 0.94 L. Finally, we evaluated twelve dose levels from 2-14 mg/kg/day with dosing intervals of 12 or 24 hours in premature infants using trial simulation. Based on the literature, the therapeutic target was determined to be a trough level of 8-20 mg/L and a peak level ≤ 40 mg/L at steady state. A dosing interval of 24 hours was shown to be successful for the proposed target concentrations in all simulated groups. With the proposed dosing regimens, the predetermined target was well attained and the simulated median trough plasma concentrations were between 8 and 20 mg/L throughout the treatment period. The dose-optimization simulations based on the developed PopPK model may provide improved therapeutic benefit while allowing clinicians to compare various dosing regimens and bridge the plasma caffeine levels with response at different PCAs and different body weights.

In summary, the general theme of my dissertation research was to investigate different approaches to overcome the unique pharmacokinetic challenges in clinical studies with premature neonates and infants. A rapid, accurate, sensitive, and reliable LC-MS/MS method to quantify five drugs frequently used in the pharmacotherapy of premature infants simultaneously was developed using a small volume of plasma. This new bioanalytical assay could facilitate an efficient use of limited blood samples in premature infants in the future. In addition, an approach was developed using population pharmacokinetic simulations to determine sample size for PopPK studies in premature neonates with the consideration of changes in birth weight, body weight and PNA. Lastly, a PopPK model was developed for caffeine in premature infants with the statistically significant relationships of clearance with weight, clearance with PCA, clearance with low GA factor, volume of distribution with weight and volume of distribution with low GA factor. Proper dosing regimens can be determined rapidly to reach the therapeutic target concentrations based on the PopPK characteristics of caffeine. Together with the LC-MS/MS bioanalytical assay, population-based modeling and simulation are highly useful in supporting clinical pharmacokinetic studies in premature neonates and infants.

With the legislative incentives and requirements, population-based modeling and simulation approaches as well as sensitive analytical assays allowing for pharmacokinetic sample quantification from very small volume blood samples, it is hoped that further information on the influence of developmental changes on pediatric pharmacokinetics will be gathered to improve pediatric drug labeling so that rationale and scientifically based dosing strategies can be developed for a safe and effective pharmacotherapy in pediatric patients.

LIST OF REFERENCES

1. U.S. Food and Drug Administration, *Comparison of NMEs approved in 2010 to previous years*, accessed February 20, 2012, from http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/DrugandBiologicApprovalReports/ucm242674.htm.

2. Lindell-Osuagwu, L., et al., *Off-label and unlicensed drug prescribing in three paediatric wards in Finland and review of the international literature.* J Clin Pharm Ther, 2009. **34**(3): p. 277-87.

3. Bazzano, A.T., et al., *Off-label prescribing to children in the United States outpatient setting.* Acad Pediatr, 2009. **9**(2): p. 81-8.

4. Winterfeld, U., et al., *Off-label use of psychotropic medications in pediatric wards: a prospective study.* Arch Pediatr, 2009. **16**(9): p. 1252-60.

5. Poole, S.G. and M.J. Dooley, *Off-label prescribing in oncology.* Support Care Cancer, 2004. **12**(5): p. 302-5.

6. 't Jong, G.W., et al., *Unlicensed and off-label prescription of respiratory drugs to children.* Eur Respir J, 2004. **23**(2): p. 310-3.

7. Bajcetic, M., et al., *Off label and unlicensed drugs use in paediatric cardiology.* Eur J Clin Pharmacol, 2005. **61**(10): p. 775-9.

8. Yoon, E.Y., et al., *Off-label utilization of antihypertensive medications in children.* Ambul Pediatr, 2007. **7**(4): p. 299-303.

9. Ufer, M., et al., *Widespread off-label prescribing of topical but not systemic drugs for 350,000 paediatric outpatients in Stockholm.* Eur J Clin Pharmacol, 2003. **58**(11): p. 779-83.

10. McIntyre, J., et al., *Unlicensed and off label prescribing of drugs in general practice.* Arch Dis Child, 2000. **83**(6): p. 498-501.

11. U.S. Food and Drug Administration, *General Considerations for Pediatric Pharmacokinetic Studies for Drugs and Biological Products — Draft Guidance.* 1998, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), FDA, Rockville, MD.

12. De Cock, R.F., et al., *The role of population PK-PD modelling in paediatric clinical research.* Eur J Clin Pharmacol, 2011. **67 Suppl 1**: p. 5-16.

13. Horen, B., J.L. Montastruc, and M. Lapeyre-Mestre, *Adverse drug reactions and off-label drug use in paediatric outpatients.* Br J Clin Pharmacol, 2002. **54**(6): p. 665-70.

14. Turner, S., et al., *Adverse drug reactions to unlicensed and off-label drugs on paediatric wards: a prospective study.* Acta Paediatr, 1999. **88**(9): p. 965-8.

15. Giacoia, G.P. and D.R. Mattison, *Newborns and drug studies: the NICHD/FDA newborn drug development initiative.* Clin Ther, 2005. **27**(6): p. 796-813.

16. Mulhall, A., J. de Louvois, and R. Hurley, *Chloramphenicol toxicity in neonates: its incidence and prevention.* Br Med J (Clin Res Ed), 1983. **287**(6403): p. 1424-7.

17. Sutherland, J.M., *Fatal cardiovascular collapse of infants receiving large amounts of chloramphenicol.* AMA J Dis Child, 1959. **97**(6): p. 761-7.

18. Johnson, T.N., *The development of drug metabolising enzymes and their influence on the susceptibility to adverse drug reactions in children.* Toxicology, 2003. **192**(1): p. 37-48.

19. Billings, R.J., R.J. Berkowitz, and G. Watson, *Teeth.* Pediatrics, 2004. **113**(4 Suppl): p. 1120-7.

20. Committee on Drugs, *Guidelines for the ethical conduct of studies to evaluate drugs in pediatric populations. Committee on Drugs, American Academy of Pediatrics.* Pediatrics, 1995. **95**(2): p. 286-94.

21. Rodriguez, W.J., R. Roberts, and D. Murphy, *Current regulatory policies regarding pediatric indications and exclusivity.* J Pediatr Gastroenterol Nutr, 2003. **37 Suppl 1**: p. S40-5.

22. Holford, N.H. and L.B. Sheiner, *Kinetics of pharmacologic response.* Pharmacol Ther, 1982. **16**(2): p. 143-66.

23. Girgis, I.G., et al., *Pharmacokinetic-pharmacodynamic assessment of topiramate dosing regimens for children with epilepsy 2 to <10 years of age.* Epilepsia, 2010. **51**(10): p. 1954-62.

24. Laer, S., et al., *Development of a safe and effective pediatric dosing regimen for sotalol based on population pharmacokinetics and pharmacodynamics in children with supraventricular tachycardia.* J Am Coll Cardiol, 2005. **46**(7): p. 1322-30.

25. Anderson, G.D., *Developmental pharmacokinetics.* Semin Pediatr Neurol, 2010. **17**(4): p. 208-13.

26. Koren, G., *Therapeutic drug monitoring principles in the neonate. National Academy of CLinical Biochemistry.* Clin Chem, 1997. **43**(1): p. 222-7.

27. Simons, F.E., H. Rigatto, and K.J. Simons, *Pharmacokinetics of theophylline in neonates.* Semin Perinatol, 1981. **5**(4): p. 337-45.

28. Hines, R.N. and D.G. McCarver, *The ontogeny of human drug-metabolizing enzymes: phase I oxidative enzymes.* J Pharmacol Exp Ther, 2002. **300**(2): p. 355-60.

29. Lacroix, D., et al., *Expression of CYP3A in the human liver--evidence that the shift between CYP3A7 and CYP3A4 occurs immediately after birth.* Eur J Biochem, 1997. **247**(2): p. 625-34.

30. Kitada, M., et al., *P-450 HFLa, a form of cytochrome P-450 purified from human fetal livers, is the 16 alpha-hydroxylase of dehydroepiandrosterone 3-sulfate.* J Biol Chem, 1987. **262**(28): p. 13534-7.

31. Anderson, G.D. and A.M. Lynn, *Optimizing pediatric dosing: a developmental pharmacologic approach.* Pharmacotherapy, 2009. **29**(6): p. 680-90.

32. Goldstein, J.A., *Clinical relevance of genetic polymorphisms in the human CYP2C subfamily.* Br J Clin Pharmacol, 2001. **52**(4): p. 349-55.

33. Koukouritaki, S.B., et al., *Developmental expression of human hepatic CYP2C9 and CYP2C19.* J Pharmacol Exp Ther, 2004. **308**(3): p. 965-74.

34. Treluyer, J.M., et al., *Expression of CYP2D6 in developing human liver.* Eur J Biochem, 1991. **202**(2): p. 583-8.

35. Jacqz-Aigrain, E. and T. Cresteil, *Cytochrome P450-dependent metabolism of dextromethorphan: fetal and adult studies.* Dev Pharmacol Ther, 1992. **18**(3-4): p. 161-8.

36. Tateishi, T., et al., *A comparison of hepatic cytochrome P450 protein expression between infancy and postinfancy.* Life Sci, 1997. **61**(26): p. 2567-74.

37. Sonnier, M. and T. Cresteil, *Delayed ontogenesis of CYP1A2 in the human liver.* Eur J Biochem, 1998. **251**(3): p. 893-8.

38. McCarver, D.G. and R.N. Hines, *The ontogeny of human drug-metabolizing enzymes: phase II conjugation enzymes and regulatory mechanisms.* J Pharmacol Exp Ther, 2002. **300**(2): p. 361-6.

39. Anderson, B.J. and N.H. Holford, *Mechanism-based concepts of size and maturity in pharmacokinetics.* Annu Rev Pharmacol Toxicol, 2008. **48**: p. 303-32.

40. Borst, P. and R.O. Elferink, *Mammalian ABC transporters in health and disease.* Annu Rev Biochem, 2002. **71**: p. 537-92.

41. Minton, N.A. and P.G. Smith, *Loperamide toxicity in a child after a single dose.* Br Med J (Clin Res Ed), 1987. **294**(6584): p. 1383.

42. Weaver, L.T., S.W. Richmond, and R. Nelson, *Loperamide toxicity in severe protracted diarrhoea.* Arch Dis Child, 1983. **58**(7): p. 568-9.

43. Daood, M., et al., *ABC transporter (P-gp/ABCB1, MRP1/ABCC1, BCRP/ABCG2) expression in the developing human CNS.* Neuropediatrics, 2008. **39**(4): p. 211-8.

44. Suzuki, H. and Y. Sugiyama, *Transporters for bile acids and organic anions.* Pharm Biotechnol, 1999. **12**: p. 387-439.

45. Bor, O., et al., *Ceftriaxone-associated biliary sludge and pseudocholelithiasis during childhood: a prospective study.* Pediatr Int, 2004. **46**(3): p. 322-4.

46. de Moor, R.A., A.C. Egberts, and C.H. Schroder, *Ceftriaxone-associated nephrolithiasis and biliary pseudolithiasis.* Eur J Pediatr, 1999. **158**(12): p. 975-7.

47. Prince, J.S. and M.O. Senac, Jr., *Ceftriaxone-associated nephrolithiasis and biliary pseudolithiasis in a child.* Pediatr Radiol, 2003. **33**(9): p. 648-51.

48. Kearns, G.L., et al., *Developmental pharmacology--drug disposition, action, and therapy in infants and children.* N Engl J Med, 2003. **349**(12): p. 1157-67.

49. Leake, R.D. and C.W. Trygstad, *Glomerular filtration rate during the period of adaptation to extrauterine life.* Pediatr Res, 1977. **11**(9 Pt 1): p. 959-62.

50. Daniel, H. and G. Kottra, *The proton oligopeptide cotransporter family SLC15 in physiology and pharmacology.* Pflugers Arch, 2004. **447**(5): p. 610-8.

51. U.S. Food and Drug Administration, *Exposure-Response Relationships — Study Design, Data Analysis, and Regulatory Applications.* 2003, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research] (CBER), FDA, Rockville, MD.

52. Tran, A., et al., *Pharmacokinetic-pharmacodynamic study of oral lansoprazole in children.* Clin Pharmacol Ther, 2002. **71**(5): p. 359-67.

53. Willis, C., C.E. Staatz, and S.E. Tett, *Bayesian forecasting and prediction of tacrolimus concentrations in pediatric liver and adult renal transplant recipients.* Ther Drug Monit, 2003. **25**(2): p. 158-66.

54. Marshall, J.D. and G.L. Kearns, *Developmental pharmacodynamics of cyclosporine.* Clin Pharmacol Ther, 1999. **66**(1): p. 66-75.

55. Saldien, V., K.M. Vermeyen, and F.L. Wuyts, *Target-controlled infusion of rocuronium in infants, children, and adults: a comparison of the pharmacokinetic and pharmacodynamic relationship.* Anesth Analg, 2003. **97**(1): p. 44-9, table of contents.

56. Takahashi, H., et al., *Developmental changes in pharmacokinetics and pharmacodynamics of warfarin enantiomers in Japanese children.* Clin Pharmacol Ther, 2000. **68**(5): p. 541-55.

57. Zajicek, A., *The National Institutes of Health and the Best Pharmaceuticals for Children Act.* Paediatr Drugs, 2009. **11**(1): p. 45-7.

58. Mathews, T.J., et al., *Annual summary of vital statistics: 2008.* Pediatrics. **127**(1): p. 146-57.

59. Partners Human Research Committee, *Blood Sample Guidelines,* accessed March 16, 2012, from http://healthcare.partners.org/phsirb/bldsamp.htm.

60. ad hoc group chaired by the European Commission, *Ethical considerations for clinical trials on medicinal products conducted with the paediatric population*. 2008.

61. Abdel-Hamid, M.E. and D. Sharma, *Simultaneous Quantification of Doxorubicin,Lorazepam, Metoclopramide, Ondansetron,and Ranitidine in Mixtures by Liquid Chromatography-Tandem Mass Spectrometry.* Journal of Liquid Chromatography & Related Technologies, 2005. **Volume 27**(Issue 4): p. 641-660.

62. Bardin, S., et al., *Determination of free levels of phenytoin in human plasma by liquid chromatography/tandem mass spectrometry.* J Pharm Biomed Anal, 2000. **23**(2-3): p. 573-9.

63. Gomez, M.J., et al., *Determination of pharmaceuticals of various therapeutic classes by solid-phase extraction and liquid chromatography-tandem mass spectrometry analysis in hospital effluent wastewaters.* J Chromatogr A, 2006. **1114**(2): p. 224-33.

64. Hori, Y., et al., *Method for screening and quantitative determination of serum levels of salicylic Acid, acetaminophen, theophylline, phenobarbital, bromvalerylurea, pentobarbital, and amobarbital using liquid chromatography/electrospray mass spectrometry.* Biol Pharm Bull, 2006. **29**(1): p. 7-13.

65. Jacob, P., 3rd, et al., *Determination of ephedra alkaloid and caffeine concentrations in dietary supplements and biological fluids.* J Anal Toxicol, 2004. **28**(3): p. 152-9.

66. Thomas, J.B., et al., *Determination of caffeine and caffeine-related metabolites in ephedra-containing standard reference materials using liquid chromatography with absorbance detection and tandem mass spectrometry.* J AOAC Int, 2007. **90**(4): p. 934-40.

67. Yin, O.Q., et al., *Rapid determination of five probe drugs and their metabolites in human plasma and urine by liquid chromatography/tandem mass spectrometry: application to cytochrome P450 phenotyping studies.* Rapid Commun Mass Spectrom, 2004. **18**(23): p. 2921-33.

68. Laer, S., J.S. Barrett, and B. Meibohm, *The in silico child: using simulation to guide pediatric drug development and manage pediatric pharmacotherapy.* J Clin Pharmacol, 2009. **49**(8): p. 889-904.

69. U.S. Food and Drug Administration, *Challenges and Opportunities Report, Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products - March 2004*. 2004.

70. Sheiner, L.B., B. Rosenberg, and V.V. Marathe, *Estimation of population characteristics of pharmacokinetic parameters from routine clinical data.* J Pharmacokinet Biopharm, 1977. **5**(5): p. 445-79.

71. U.S. Food and Drug Administration, *Population Pharmacokinetics Guidance for Industry,* 1999, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), FDA, Rockville, MD.

72. Sheiner, L.B. and T.M. Ludden, *Population pharmacokinetics/dynamics.* Annu Rev Pharmacol Toxicol, 1992. **32**: p. 185-209.

73. Meibohm, B., et al., *Population pharmacokinetic studies in pediatrics: issues in design and analysis.* Aaps J, 2005. **7**(2): p. E475-87.

74. Wahlby, U., E.N. Jonsson, and M.O. Karlsson, *Comparison of stepwise covariate model building strategies in population pharmacokinetic-pharmacodynamic analysis.* AAPS PharmSci, 2002. **4**(4): p. E27.

75. Jonsson, E.N. and M.O. Karlsson, *Automated covariate model building within NONMEM.* Pharm Res, 1998. **15**(9): p. 1463-8.

76. Barrett, J.S., *Pharmacokinetics in Drug Discovery and Development*. Chapter 15 Population Pharmacokinetics, ed. R.D. Schoenwald. 2002: CRC Press

77. Meibohm, B. and H. Derendorf, *Pharmacokinetic/pharmacodynamic studies in drug product development.* J Pharm Sci, 2002. **91**(1): p. 18-31.

78. Gobburu, J.V., *Pharmacometrics 2020.* J Clin Pharmacol, 2010. **50**(9 Suppl): p. 151S-157S.

79. Lee, J.Y., et al., *Impact of pharmacometric analyses on new drug approval and labelling decisions: a review of 198 submissions between 2000 and 2008.* Clin Pharmacokinet, 2011. **50**(10): p. 627-35.

80. Buck, M.L., *Impact of new regulations for pediatric labeling by the Food and Drug Administration.* Pediatr Nurs, 2000. **26**(1): p. 95-6.

81. Roberts, R., et al., *Pediatric drug labeling: improving the safety and efficacy of pediatric therapies.* Jama, 2003. **290**(7): p. 905-11.

82. Lo, Y.L., et al., *Population pharmacokinetics of vancomycin in premature Malaysian neonates: identification of predictors for dosing determination.* Antimicrob Agents Chemother, 2010. **54**(6): p. 2626-32.

83. Anderson, B.J., et al., *Vancomycin pharmacokinetics in preterm neonates and the prediction of adult clearance.* Br J Clin Pharmacol, 2007. **63**(1): p. 75-84.

84. Marques-Minana, M.R., A. Saadeddin, and J.E. Peris, *Population pharmacokinetic analysis of vancomycin in neonates. A new proposal of initial dosage guideline.* Br J Clin Pharmacol, 2010. **70**(5): p. 713-20.

85. Al Za'abi, M., et al., *Application of routine monitoring data for determination of the population pharmacokinetics and enteral bioavailability of phenytoin in neonates and infants with seizures.* Ther Drug Monit, 2006. **28**(6): p. 793-9.

86. Lee, T.C., et al., *Population pharmacokinetic modeling in very premature infants receiving midazolam during mechanical ventilation: midazolam neonatal pharmacokinetics.* Anesthesiology, 1999. **90**(2): p. 451-7.

87. Touw, D.J., E.M. Westerman, and A.J. Sprij, *Therapeutic drug monitoring of aminoglycosides in neonates.* Clin Pharmacokinet, 2009. **48**(2): p. 71-88.

88. De Ridder, F., *Predicting the outcome of phase III trials using phase II data: a case study of clinical trial simulation in late stage drug development.* Basic Clin Pharmacol Toxicol, 2005. **96**(3): p. 235-41.

89. Bonate, P.L., *Clinical trial simulation in drug development.* Pharm Res, 2000. **17**(3): p. 252-6.

90. Mouksassi, M.S., et al., *Clinical trial simulations in pediatric patients using realistic covariates: application to teduglutide, a glucagon-like peptide-2 analog in neonates and infants with short-bowel syndrome.* Clin Pharmacol Ther, 2009. **86**(6): p. 667-71.

91. Bhattaram, V.A., et al., *Impact of pharmacometric reviews on new drug approval and labeling decisions--a survey of 31 new drug applications submitted between 2005 and 2006.* Clin Pharmacol Ther, 2007. **81**(2): p. 213-21.

92. Bhattaram, V.A., et al., *Impact of pharmacometrics on drug approval and labeling decisions: a survey of 42 new drug applications.* Aaps J, 2005. **7**(3): p. E503-12.

93. Wang, Y., et al., *Leveraging prior quantitative knowledge to guide drug development decisions and regulatory science recommendations: impact of FDA pharmacometrics during 2004-2006.* J Clin Pharmacol, 2008. **48**(2): p. 146-56.

94. Li, F., et al., *Pharmacometrics-based dose selection of levofloxacin as a treatment for postexposure inhalational anthrax in children.* Antimicrob Agents Chemother, 2010. **54**(1): p. 375-9.

95. Zhang, Y., et al., *A tandem mass spectrometry assay for the simultaneous determination of acetaminophen, caffeine, phenytoin, ranitidine, and theophylline in small volume pediatric plasma specimens.* Clin Chim Acta, 2008. **398**(1-2): p. 105-12.

96. Alcorn, J. and P.J. McNamara, *Pharmacokinetics in the newborn.* Adv Drug Deliv Rev, 2003. **55**(5): p. 667-86.

97. Christensen, M.L., R.A. Helms, and R.W. Chesney, *Is pediatric labeling really necessary?* Pediatrics, 1999. **104**(3 Pt 2): p. 593-7.

98. Arana, A., N.S. Morton, and T.G. Hansen, *Treatment with paracetamol in infants.* Acta Anaesthesiol Scand, 2001. **45**(1): p. 20-9.

99. Aranda, J.V., et al., *Efficacy of caffeine in treatment of apnea in the low-birth-weight infant.* J Pediatr, 1977. **90**(3): p. 467-72.

100. Aranda, J.V., D. Grondin, and B.I. Sasyniuk, *Pharmacologic considerations in the therapy of neonatal apnea.* Pediatr Clin North Am, 1981. **28**(1): p. 113-33.

101. Erenberg, A., et al., *Caffeine citrate for the treatment of apnea of prematurity: a double-blind, placebo-controlled study.* Pharmacotherapy, 2000. **20**(6): p. 644-52.

102. Kelly, E.J., et al., *The effect of intravenous ranitidine on the intragastric pH of preterm infants receiving dexamethasone.* Arch Dis Child, 1993. **69**(1 Spec No): p. 37-9.

103. Frey, O.R., A.I. von Brenndorff, and W. Probst, *Comparison of phenytoin serum concentrations in premature neonates following intravenous and oral administration.* Ann Pharmacother, 1998. **32**(3): p. 300-3.

104. Meibohm, B., *LC-MS/MS Combined with POPPK, Challenges of Pharmacokinetic Evaluations in Pediatric Populations.* G.I.T. Laboratory Journal, 2006: p. 36-37.

105. U.S. Food and Drug Administration, *Guidance for Industry, bioanalytical Method Validation*. 2001, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM), FDA, Rockville, MD.

106. Li, S., et al., *Simultaneous determination of ten antiarrhythic drugs and a metabolite in human plasma by liquid chromatography--tandem mass spectrometry.* J Chromatogr B Analyt Technol Biomed Life Sci, 2007. **847**(2): p. 174-81.

107. Matuszewski, B.K., M.L. Constanzer, and C.M. Chavez-Eng, *Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS.* Anal Chem, 2003. **75**(13): p. 3019-30.

108. Zhang, G., A.V. Terry, Jr., and M.G. Bartlett, *Liquid chromatography/tandem mass spectrometry method for the simultaneous determination of olanzapine, risperidone, 9-hydroxyrisperidone, clozapine, haloperidol and ziprasidone in rat plasma.* Rapid Commun Mass Spectrom, 2007. **21**(6): p. 920-8.

109. Jiao, J., et al., *Optimization of triple quadrupole mass spectrometer for quantitation of trace degradants of pharmaceutical compounds.* International Journal of Mass Spectrometry, 2002. **216**: p. 209–218.

110. Nahata, M.C., et al., *Acetaminophen accumulation in pediatric patients after repeated therapeutic doses.* Eur J Clin Pharmacol, 1984. **27**(1): p. 57-9.

111. van Lingen, R.A., et al., *Multiple-dose pharmacokinetics of rectally administered acetaminophen in term infants.* Clin Pharmacol Ther, 1999. **66**(5): p. 509-15.

112. Lowry, J.A., et al., *Theophylline toxicokinetics in premature newborns.* Arch Pediatr Adolesc Med, 2001. **155**(8): p. 934-9.

113. Pesce, A.J., M. Rashkin, and U. Kotagal, *Standards of laboratory practice: theophylline and caffeine monitoring. National Academy of Clinical Biochemistry.* Clin Chem, 1998. **44**(5): p. 1124-8.

114. Yaffe, S.J. and J.V. Aranda, *Neonatal and Pediatric Pharmacology: Therapeutic Principles in Practice* Third ed. 2005. 505.

115. Fontana, M., et al., *Ranitidine pharmacokinetics in newborn infants.* Arch Dis Child, 1993. **68**(5 Spec No): p. 602-3.

116. Rosenthal, M. and P.W. Miller, *Ranitidine in the newborn.* Arch Dis Child, 1988. **63**(1): p. 88-9.

117. Wells, T.G., et al., *Pharmacokinetics and pharmacodynamics of ranitidine in neonates treated with extracorporeal membrane oxygenation.* J Clin Pharmacol, 1998. **38**(5): p. 402-7.

118. Thomson, A.H. and B. Whiting, *Bayesian parameter estimation and population pharmacokinetics.* Clin Pharmacokinet, 1992. **22**(6): p. 447-67.

119. Yukawa, M., et al., *Population pharmacokinetic investigation of digoxin in Japanese infants and young children.* J Clin Pharmacol, 2011. **51**(6): p. 857-63.

120. Booth, B.P. and J.V. Gobburu, *Considerations in analyzing single-trough concentrations using mixed-effects modeling.* J Clin Pharmacol, 2003. **43**(12): p. 1307-15.

121. Aarons, L., *Population pharmacokinetics: theory and practice.* Br J Clin Pharmacol, 1991. **32**(6): p. 669-70.

122. Fadiran, E.O., C.D. Jones, and E.I. Ette, *Designing population pharmacokinetic studies: performance of mixed designs.* Eur J Drug Metab Pharmacokinet, 2000. **25**(3-4): p. 231-9.

123. Kang, D., J.B. Schwartz, and D. Verotta, *A sample size computation method for non-linear mixed effects models with applications to pharmacokinetics models.* Stat Med, 2004. **23**(16): p. 2551-66.

124. Kang, D., J.B. Schwartz, and D. Verotta, *Sample size computations for PK/PD population models.* J Pharmacokinet Pharmacodyn, 2005. **32**(5-6): p. 685-701.

125. White, D.B., et al., *Evaluation of hypothesis testing for comparing two populations using NONMEM analysis.* J Pharmacokinet Biopharm, 1992. **20**(3): p. 295-313.

126. Lee, P.I., *Design and power of a population pharmacokinetic study.* Pharm Res, 2001. **18**(1): p. 75-82.

127. Ogungbenro, K. and L. Aarons, *How many subjects are necessary for population pharmacokinetic experiments? Confidence interval approach.* Eur J Clin Pharmacol, 2008. **64**(7): p. 705-13.

128. Anderson, B.J., et al., *Acetaminophen developmental pharmacokinetics in premature neonates and infants: a pooled population analysis.* Anesthesiology, 2002. **96**(6): p. 1336-45.

129. Mulla, H., et al., *Population pharmacokinetics of theophylline during paediatric extracorporeal membrane oxygenation.* Br J Clin Pharmacol, 2003. **55**(1): p. 23-31.

130. Aranda, J.V., et al., *Maturation of caffeine elimination in infancy.* Arch Dis Child, 1979. **54**(12): p. 946-9.

131. Lee, T.C., et al., *Theophylline population pharmacokinetics from routine monitoring data in very premature infants with apnoea.* Br J Clin Pharmacol, 1996. **41**(3): p. 191-200.

132. Ehrenkranz, R.A., et al., *Longitudinal growth of hospitalized very low birth weight infants.* Pediatrics, 1999. **104**(2 Pt 1): p. 280-9.

133. Rosner, B., *Fundamentals of Biostatistics*. 7th ed. Chapter 3 Probability. 2011: Brooks/Cole, Cengage Learning.

134. Schuirmann, D.J., *A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability.* J Pharmacokinet Biopharm, 1987. **15**(6): p. 657-80.

135. Ogungbenro, K. and L. Aarons, *Sample size/power calculations for repeated ordinal measurements in population pharmacodynamic experiments.* J Pharmacokinet Pharmacodyn, 2010. **37**(1): p. 67-83.

136. Kowalski, K.G. and M.M. Hutmacher, *Design evaluation for a population pharmacokinetic study using clinical trial simulations: a case study.* Stat Med, 2001. **20**(1): p. 75-91.

137. Beal, S.L., *Sample size determination for confidence intervals on the population mean and on the difference between two population means.* Biometrics, 1989. **45**(3): p. 969-77.

138. Ette, E.I., H. Sun, and T.M. Ludden, *Balanced designs in longitudinal population pharmacokinetic studies.* J Clin Pharmacol, 1998. **38**(5): p. 417-23.

139. Colucci, P., et al., *Performance of different population pharmacokinetic algorithms.* Ther Drug Monit. **33**(5): p. 583-91.

140. Yang, S. and M. Beerahee, *Power estimation using a population pharmacokinetics model with optimal design by clinical trial simulations: application in pharmacokinetic drug-drug interaction studies.* Eur J Clin Pharmacol. **67**(3): p. 225-33.

141. Panetta, J.C., et al., *The importance of pharmacokinetic limited sampling models for childhood cancer drug development.* Clin Cancer Res, 2003. **9**(14): p. 5068-77.

142. Retout, S., S. Duffull, and F. Mentre, *Development and implementation of the population Fisher information matrix for the evaluation of population pharmacokinetic designs.* Comput Methods Programs Biomed, 2001. **65**(2): p. 141-51.

143. Retout, S., F. Mentre, and R. Bruno, *Fisher information matrix for non-linear mixed-effects models: evaluation and application for optimal design of enoxaparin population pharmacokinetics.* Stat Med, 2002. **21**(18): p. 2623-39.

144. Schmidt, B., et al., *Caffeine therapy for apnea of prematurity.* N Engl J Med, 2006. **354**(20): p. 2112-21.

145. Poets, C.F., *Interventions for apnoea of prematurity: a personal view.* Acta Paediatr. **99**(2): p. 172-7.

146. Schmidt, B., et al., *Long-term effects of caffeine therapy for apnea of prematurity.* N Engl J Med, 2007. **357**(19): p. 1893-902.

147. Papageorgiou, A., Pelausa, E. ,Kovacs, L., *Avery's Neonatology: Pathophysiology & Management of the Newborn*. Part IV Chapter 25 The Extremely Low-Birth-Weight Infant, ed. M.G. MacDonald, Seshia, Mary M.K., Mullett, Martha D. 2005, : Lippincott Williams & Wilkins.

148. Aranda, J.V., et al., *Caffeine impact on neonatal morbidities.* J Matern Fetal Neonatal Med, 2010 **23 Suppl 3**: p. 20-3.

149. Jonsson, E.N. and M.O. Karlsson, *Xpose--an S-PLUS based population pharmacokinetic/pharmacodynamic model building aid for NONMEM.* Comput Methods Programs Biomed, 1999. **58**(1): p. 51-64.

150. Wilkins, J.J., *NONMEMory: a run management tool for NONMEM.* Comput Methods Programs Biomed, 2005. **78**(3): p. 259-67.

151. Falcao, A.C., et al., *Population pharmacokinetics of caffeine in premature neonates.* Eur J Clin Pharmacol, 1997. **52**(3): p. 211-7.

152. Charles, B.G., et al., *Caffeine citrate treatment for extremely premature infants with apnea: population pharmacokinetics, absolute bioavailability, and implications for therapeutic drug monitoring.* Ther Drug Monit, 2008. **30**(6): p. 709-16.

153. Aranda, J.V., et al., *Pharmacokinetic profile of caffeine in the premature newborn infant with apnea.* J Pediatr, 1979. **94**(4): p. 663-8.

154. Newton, R., et al., *Plasma and salivary pharmacokinetics of caffeine in man.* Eur J Clin Pharmacol, 1981. **21**(1): p. 45-52.

155. Karlsson, M.O. and R.M. Savic, *Diagnosing model diagnostics.* Clin Pharmacol Ther, 2007. **82**(1): p. 17-20.

156. Bonate, P.L., *The effect of collinearity on parameter estimates in nonlinear mixed effect models.* Pharm Res, 1999. **16**(5): p. 709-17.

157. Savic, R.M. and M.O. Karlsson, *Importance of shrinkage in empirical bayes estimates for diagnostics: problems and solutions.* Aaps J, 2009. **11**(3): p. 558-69.

158. Aranda, J.V. and T. Turmen, *Methylxanthines in apnea of prematurity.* Clin Perinatol, 1979. **6**(1): p. 87-108.

159. West, G.B., J.H. Brown, and B.J. Enquist, *The fourth dimension of life: fractal geometry and allometric scaling of organisms.* Science, 1999. **284**(5420): p. 1677-9.

160. Le Guennec, J.C., B. Billon, and C. Pare, *Maturational changes of caffeine concentrations and disposition in infancy during maintenance therapy for apnea of prematurity: influence of gestational age, hepatic disease, and breast-feeding.* Pediatrics, 1985. **76**(5): p. 834-40.

161. Johnson, T.N., A. Rostami-Hodjegan, and G.T. Tucker, *Prediction of the clearance of eleven drugs and associated variability in neonates, infants and children.* Clin Pharmacokinet, 2006. **45**(9): p. 931-56.

162. Gorodischer, R. and M. Karplus, *Pharmacokinetic aspects of caffeine in premature infants with apnoea.* Eur J Clin Pharmacol, 1982. **22**(1): p. 47-52.

163. Thomson, A.H., S. Kerr, and S. Wright, *Population pharmacokinetics of caffeine in neonates and young infants.* Ther Drug Monit, 1996. **18**(3): p. 245-53.

164. Lee, T.C., et al., *Population pharmacokinetics of intravenous caffeine in neonates with apnea of prematurity.* Clin Pharmacol Ther, 1997. **61**(6): p. 628-40.

165. Micallef, S., et al., *Sequential updating of a new dynamic pharmacokinetic model for caffeine in premature neonates.* Clin Pharmacokinet, 2007. **46**(1): p. 59-74.

166. al-Alaiyan, S., et al., *Caffeine metabolism in premature infants.* J Clin Pharmacol, 2001. **41**(6): p. 620-7.

167. Lee, H.S., et al., *Caffeine in apnoeic Asian neonates: a sparse data analysis.* Br J Clin Pharmacol, 2002. **54**(1): p. 31-7.

168. McNamara, P.J. and J. Alcorn, *Protein binding predictions in infants.* AAPS PharmSci, 2002. **4**(1): p. E4.

169. Blanchard, J., *Protein binding of caffeine in young and elderly males.* J Pharm Sci, 1982. **71**(12): p. 1415-8.

170. Ginsberg, G., et al., *Physiologically based pharmacokinetic (PBPK) modeling of caffeine and theophylline in neonates and adults: implications for assessing children's risks from environmental agents.* J Toxicol Environ Health A, 2004. **67**(4): p. 297-329.

171. Fomon, S.J., et al., *Body composition of reference children from birth to age 10 years.* Am J Clin Nutr, 1982. **35**(5 Suppl): p. 1169-75.

172. Pons, G., et al., *Developmental changes of caffeine elimination in infancy.* Dev Pharmacol Ther, 1988. **11**(5): p. 258-64.

173. Wade, K.C., et al., *Population pharmacokinetics of fluconazole in young infants.* Antimicrob Agents Chemother, 2008. **52**(11): p. 4043-9.

174. Aldridge, A., J.V. Aranda, and A.H. Neims, *Caffeine metabolism in the newborn.* Clin Pharmacol Ther, 1979. **25**(4): p. 447-53.

175. Cornish, H.H. and A.A. Christman, *A study of the metabolism of theobromine, theophylline, and caffeine in man.* J Biol Chem, 1957. **228**(1): p. 315-23.

176. Lelo, A., et al., *Quantitative assessment of caffeine partial clearances in man.* Br J Clin Pharmacol, 1986. **22**(2): p. 183-6.

177. Miners, J.O. and D.J. Birkett, *The use of caffeine as a metabolic probe for human drug metabolizing enzymes.* Gen Pharmacol, 1996. **27**(2): p. 245-9.

178. Rost, K.L. and I. Roots, *Accelerated caffeine metabolism after omeprazole treatment is indicated by urinary metabolite ratios: coincidence with plasma clearance and breath test.* Clin Pharmacol Ther, 1994. **55**(4): p. 402-11.

179. Cazeneuve, C., et al., *Biotransformation of caffeine in human liver microsomes from foetuses, neonates, infants and adults.* Br J Clin Pharmacol, 1994. **37**(5): p. 405-12.

180. Schmidt, R.J., et al., *Caffeine, selected metabolic gene variants, and risk for neural tube defects.* Birth Defects Res A Clin Mol Teratol, 2010. **88**(7): p. 560-9.

181. Alcorn, J. and P.J. McNamara, *Ontogeny of hepatic and renal systemic clearance pathways in infants: part I.* Clin Pharmacokinet, 2002. **41**(12): p. 959-98.

182. Horning, M.G., et al., *Drug metabolism in the human neonate.* Life Sci, 1975. **16**(5): p. 651-71.

183. Carrier, O., et al., *Maturation of caffeine metabolic pathways in infancy.* Clin Pharmacol Ther, 1988. **44**(2): p. 145-51.

184. Gallini, F., et al., *Progression of renal function in preterm neonates with gestational age < or = 32 weeks.* Pediatr Nephrol, 2000. **15**(1-2): p. 119-24.

185. van den Anker, J.N., et al., *Ceftazidime pharmacokinetics in preterm infants: effects of renal function and gestational age.* Clin Pharmacol Ther, 1995. **58**(6): p. 650-9.

186. Sluncheva, B., A. Dimitrov, and L. Vakrilova, *Development of renal function in low and extremely low birth weight infants: correlation with gestational and postnatal age (Abstract Only).* Akush Ginekol (Sofiia), 2002. **41**(5): p. 20-6.

187. Van Den Anker, J.N., et al., *The effect of asphyxia on the pharmacokinetics of ceftazidime in the term newborn.* Pediatr Res, 1995. **38**(5): p. 808-11.

188. Mannan, M.A., et al., *Postnatal development of renal function in preterm and term neonates.* Mymensingh Med J, 2012. **21**(1): p. 103-8.

APPENDIX A. CHAPTER 4 SUPPLEMENTAL FIGURES

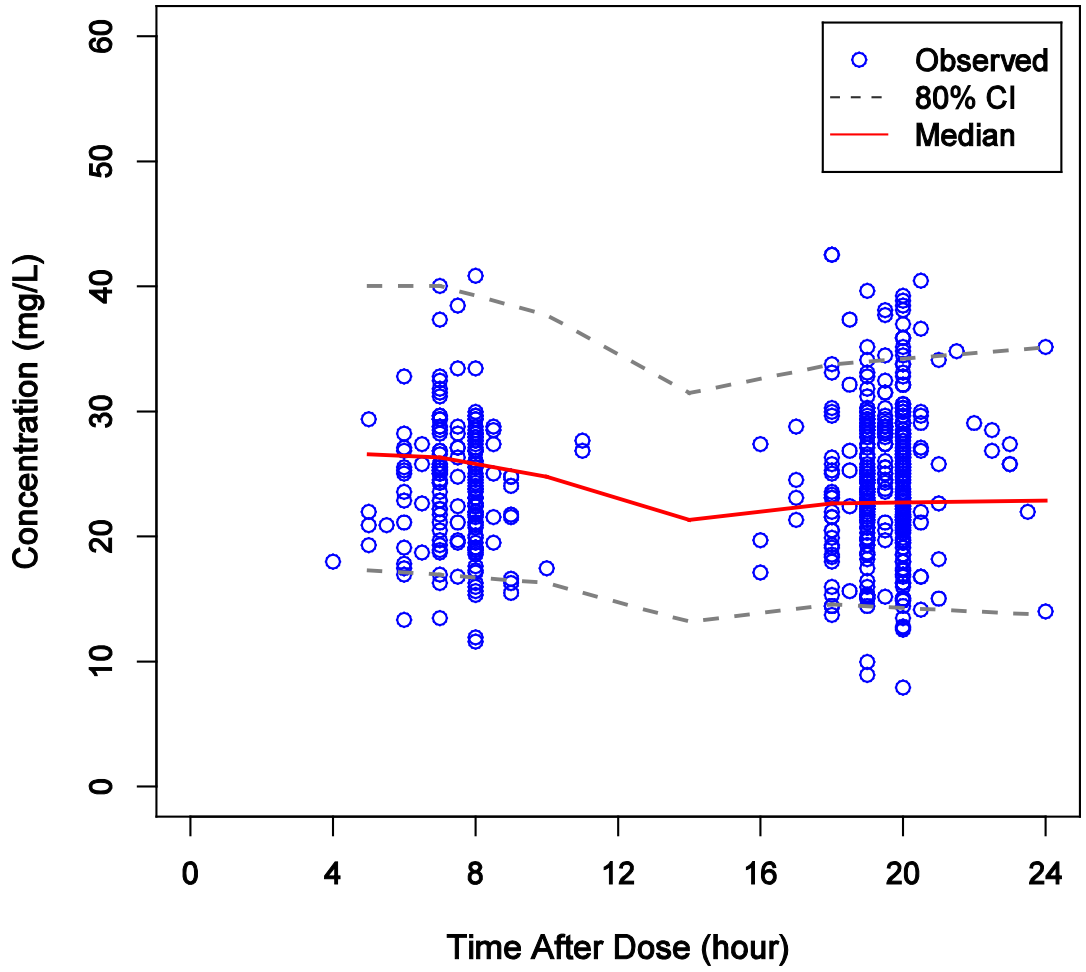


Figure A-1. Visual predictive check for the final caffeine population pharmacokinetic model

The red solid line and grey dotted lines indicate the median and 80% confidence interval of predicted concentrations determined from 500 Monte Carlo simulations with the Final Model. Open circles indicate observed caffeine concentrations in the model building dataset.

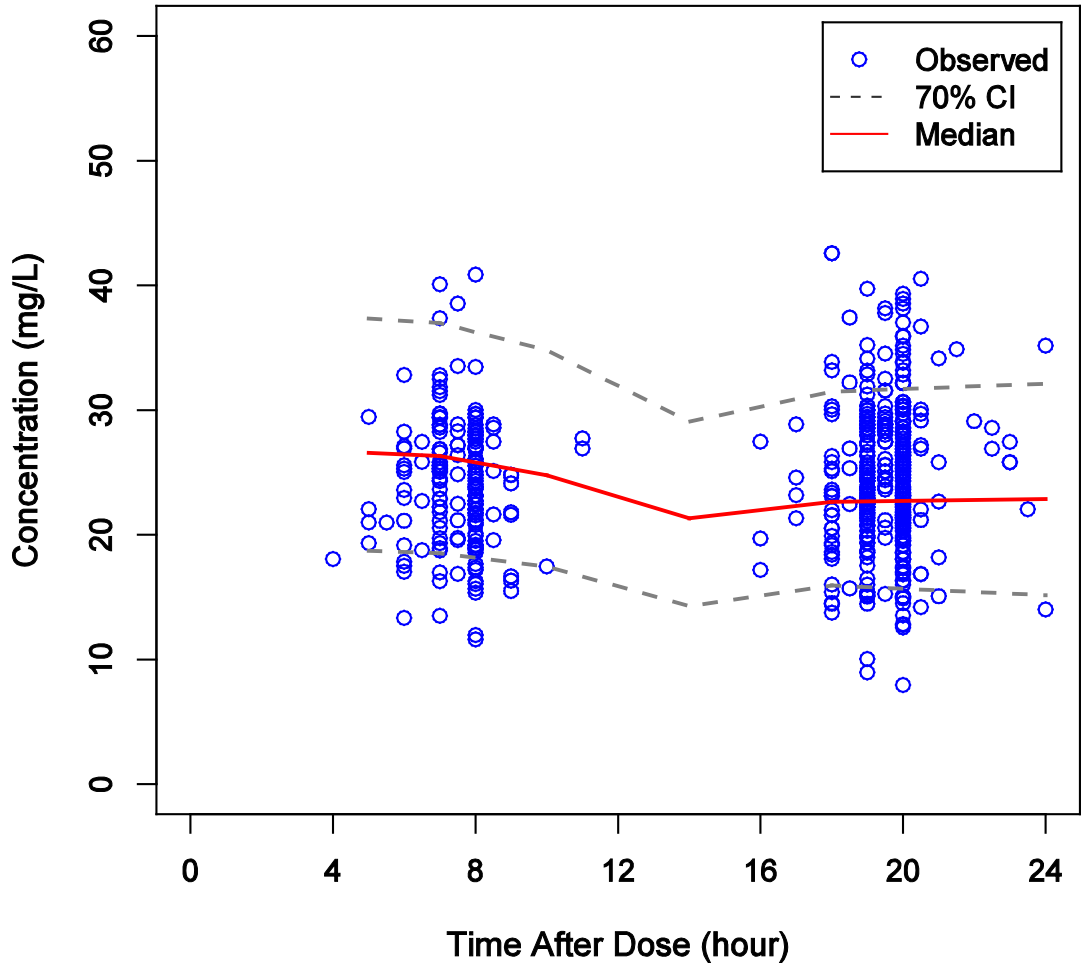


Figure A-2. Visual predictive check for the final caffeine population pharmacokinetic model

The red solid line and grey dotted lines indicate the median and 70% confidence interval of predicted concentrations determined from 500 Monte Carlo simulations with the Final Model. Open circles indicate observed caffeine concentrations in the model building dataset.

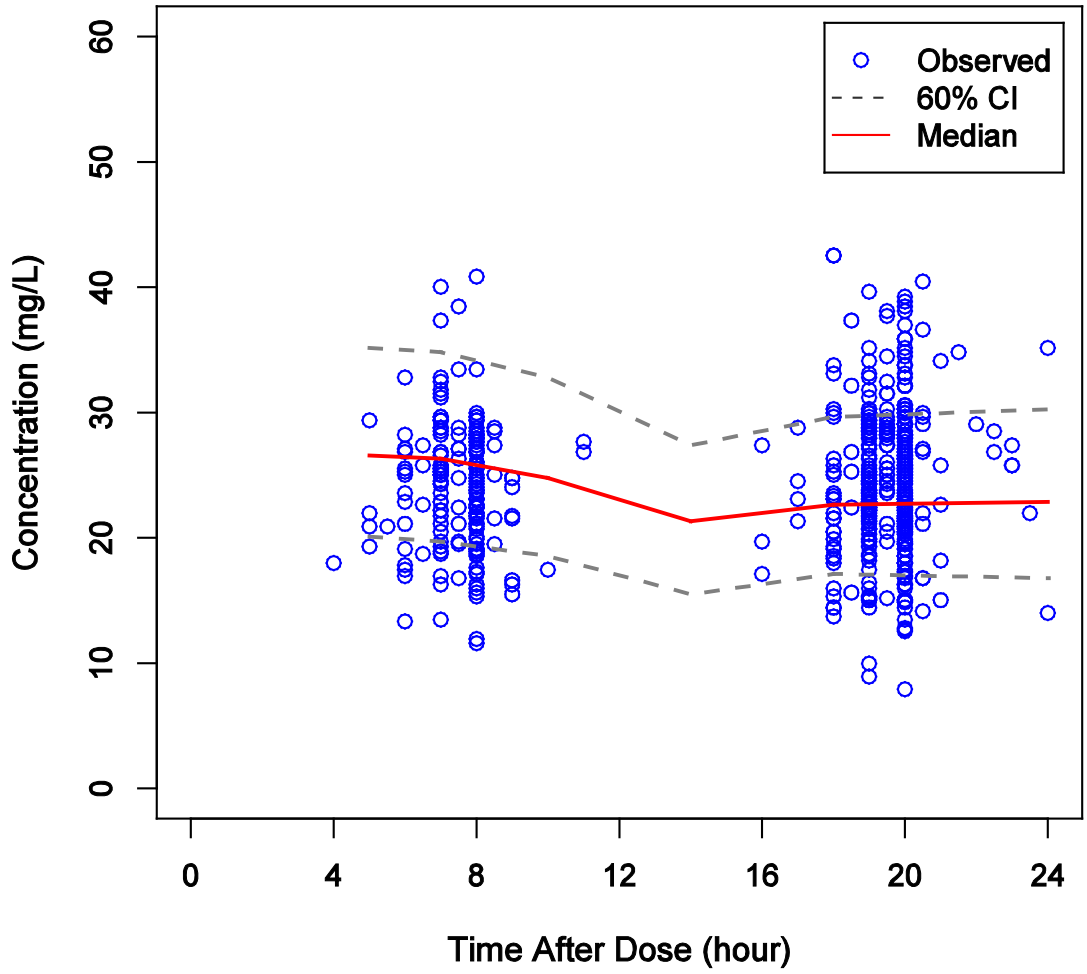


Figure A-3. Visual predictive check for the final caffeine population pharmacokinetic model

The red solid line and grey dotted lines indicate the median and 60% confidence interval of predicted concentrations determined from 500 Monte Carlo simulations with the Final Model. Open circles indicate observed caffeine concentrations in the model building dataset.

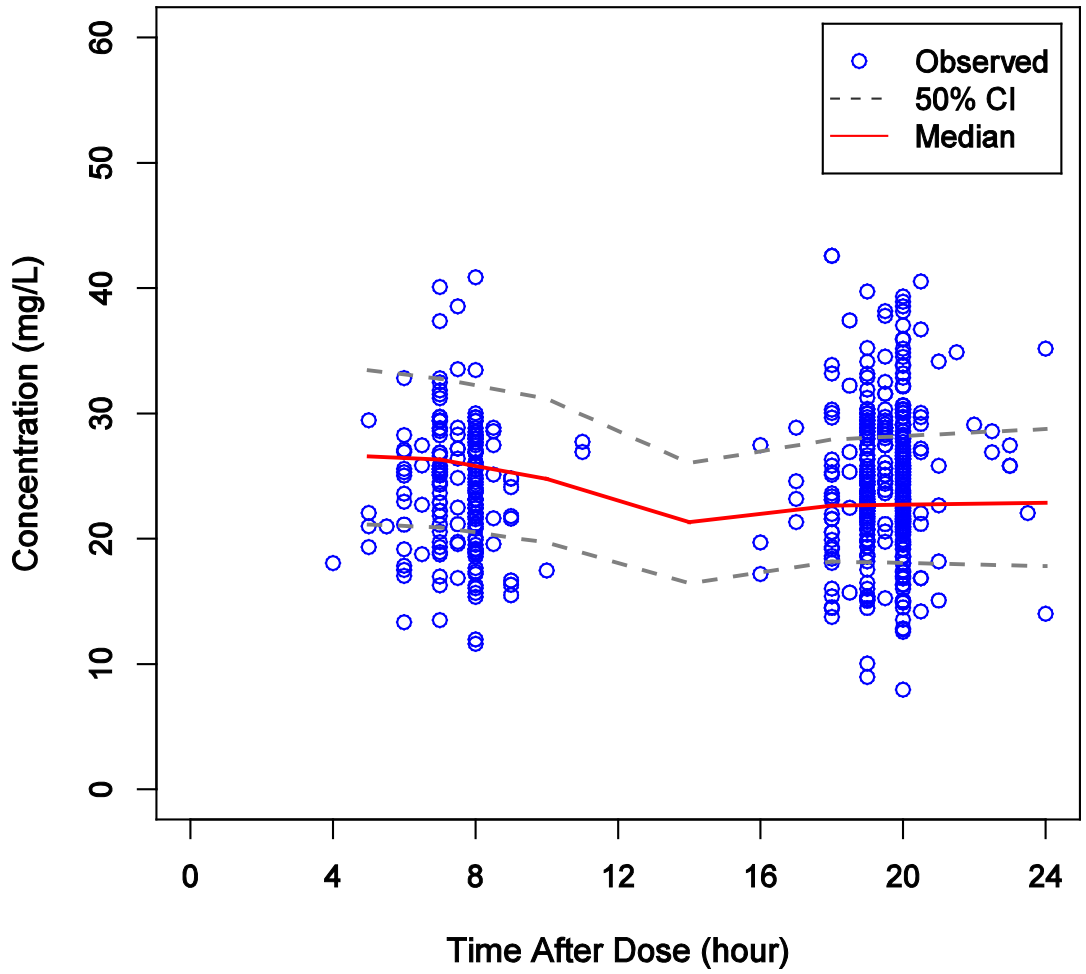


Figure A-4. Visual predictive check for the final caffeine population pharmacokinetic model

The red solid line and grey dotted lines indicate the median and 50% confidence interval of predicted concentrations determined from 500 Monte Carlo simulations with the Final Model. Open circles indicate observed caffeine concentrations in the model building dataset.

VITA

Yi Zhang, daughter of Mr. Hongrui Zhang and Ms. Jingzun Wang, was born in Tianjin, China in 1972. She graduated from Tianjin Medical University with a Doctor of Medicine degree in 1995. After graduation, she joined the department of Gerontology at Tianjin Medical University General Hospital as resident, chief resident, physician and attending physician. In 2006, she enrolled in the Ph.D. program at the University of Tennessee Health Science Center under the direction of Dr. Bernd Meibohm in the Department of Pharmaceutical Sciences. She expects to receive her Doctor of Philosophy degree in May, 2012. She is a member of the American Association of Pharmaceutical Scientists and the American College of Clinical Pharmacology.